DOCUMENT RESUME

ED 170 220 SO 011 674

AUTHOR Perlmann, Ari Joel
TITLE Education and Social Structure: Providence, R.I. 1880-1925. The Research Design and the Data Collection.
SPONS AGENCY National Inst. of Education (DHEW), Washington, D.C.
PUB DATE 20 Mar 78
GRANT NIE-G-740075
NOTE 136p.; Table on page 59 and Data Collection charts on last four pages of document may not reproduce clearly due to small print type of original

EDRS PRICE MF01/PC06 Plus Postage.
DESCRIPTORS *Academic Achievement; Census Figures; Community Study; Data; *Data Collection; Economic Research; Educational History; Educational Research; Elementary Secondary Education; Family Background; Relationship; Research Methodology; Research Problems; Research Tools; School Statistics; Social Indicators; *Social Mobility; *Social Science Research; Socioeconomic Background; Socioeconomic Status; Vocational Followup
IDENTIFIERS Rhode Island (Providence)

ABSTRACT

This interim report describes in detail the data collection procedures for a study focusing on the relation of education to social structure in Providence, Rhode Island from 1880 to 1925. The objectives of the study are to examine the following: (1) the changing distribution of schooling among various social groups; (2) the extent to which schools helped people get ahead; (3) the extent to which different ethnic groups differed in their use of the schools; (4) the educational patterns in women of different social backgrounds and the role of education in their social mobility; and (5) the extent to which the educational credentials required for a job changed as the average amount of schooling increased. The author details how he chose samples from the census and high school records, how he obtained family background information from the city directory and the tax book, information on school attainment from the census and school records, and information on the sample members' careers, male and female, from the city directory, tax books, and marriage records. Limitations, difficulties, and the possible margin of error are noted in full. The author also describes how he classified and coded occupations and concludes with a description of organizing and processing the data collection. Research findings and implications will be reported at a later date. The data collection chart used is appended. (Author/CK)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED170220

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

EDUCATION AND SOCIAL STRUCTURE:
PROVIDENCE, R.I. 1880-1925

THE RESEARCH DESIGN AND
THE DATA COLLECTION

Final Report to the National
Institute of Education
Grant Number NIE-G-740075

Ari Joel Perlmann
March 20, 1978

SO 011 674

Contents

An Explanatory Foreword i

Introduction iii-xix

I. Overview: The Design 1
II. Selecting the Samples 7
    A. The Census Samples
    B. The High School Samples
III. Obtaining Additional Information on Family Background 21
    A. Tracing Fathers to the City Directory and to the Tax Book
    B. Tracing High School Sample Members to the Census Manuscripts
IV. Obtaining Information on School Attainment 26
    A. The Limitations of Census and Survey Data for Historical Study of School Attainment
    B. A Description of the School Records
    C. Linking Sample Members to School Records
    D. The Completeness of the Records and the Accuracy of the Linkage Procedure
V. Obtaining Information on the Sample Members' Careers 73
    A. Tracing Male Sample Members Across Time to the City Directory and the Tax Books
    B. Obtaining Information on Female Sample Members from the Marriage Records
        1. Tracing the Census and High School Samples to the Marriage Records
        2. Collecting a Sample of Brides in White Collar Occupations
VI. The Classification and Coding of Occupations 79
VII. The Organization of the Data Collection and Processing 86

Notes follow p.88

An Explanatory Foreword

This is the final report on a project funded by a National Institute of Education grant (Grant number NIE-G-740075). But the research effort which that funding initiated is continuing. Therefore from an intellectual, rather than an administrative, point of view, this document should be regarded as an interim report.

The N.I.E. supported the research from July 1, 1974, to September 30, 1976 (although the grant period formally was extended until April 15, 1977, all funds were expended by the date indicated). During that period, about 60% of the data now available were collected, keypunched, verified, thoroughly cleaned, transformed in various other preliminary ways, and in general made ready for analysis. Moreover, much of the start-up effort required for the later data collection had already been undertaken for this first 60% and did not have to be repeated: studying the sources, preparing forms and directions, piloting them, creating various computer programs, and so forth.

The Center for the Study of Metropolitan Problems of the National Institute of Mental Health has supported the extension of the research.

When the period of N.I.E. funding came to a close my project officers (first Ella Kelly and then Reed Beddow) and I agreed that I should submit a final report on the N.I.E.-supported work, even if the research was continuing. I felt that a methodological paper having to do with the nature of the material collected would be the most appropriate closure to the project. I suggested this to Reed Bellow and he agreed. What follows, therefore, is primarily a detailed description of the data collection effort. That description provides, I think, a reasonably close look at the data social historians use, and the ways in which it can be handled. The research of social historians is likely to bear increasingly on studies of social stratification and on contemporary social research generally by adding a time dimension to these studies. This quality has always been claimed for historical work. But the current concerns, sources, and methods of social history are much more of a piece with studies of contemporary American society than was true in the past. So "adding a time dimension" should prove to be more than a cliché.

But the bulk of the material presented will also be new to most social historians, for it deals with the use of school records. These records constitute a virtually untouched source. And I hope that what I have to say about their value and their use will be of interest to sociologists interested in education, to historians of education, and to social historians who have not been concerned with education.

In addition I have included an introduction. This section includes a recent statement of the issues which prompted the collection of the data in the first place. This statement was used as the basis of the proposal to the N.I.M.H. and is similar to the statement of the problem first submitted to the N.I.E. But this final report does not address the issues raised in the introduction. It merely describes the material with which I expect to do so. Still, the niceties of the research methods will be more meaningful to readers if the substantive issues are understood.

The actual research findings will be reported in part in my doctoral dissertation in the Department of History at Harvard, and subsequently in expanded published form as a book. The piece included here will also constitute (in somewhat modified form) a methodological section in that work. Also some of the issues treated here will be discussed in modified form in journal articles. Copies of related work, of course, also be sent to the N.I.E. as it appears.

INTRODUCTION

The Proposal in General. This is a study of the relation of education to social structure in the American past by analyzing first, the relationship between social background and educational attainment and second, the impact of education on social mobility. It will deal mainly with primary and secondary education and focus especially on the lower part of the social hierarchy. The work will rest heavily on a linkage of information about the social origins of an individual (in particular, the occupation and property of his father, and his nativity or race) with his school records (which tell how much education he received, in what program, and how well he did academically) and with information about his own socioeconomic position in later life (occupation and property).[1]

The Research Context. Some historical studies have tried to analyze how educational institutions have contributed to the stability or change of the social structure. But they usually focused either on the evolution of school programs aimed at particular social groups or on the political struggles surrounding educational issues. Less often one finds analyses of which groups of people actually attended the schools or how their subsequent careers might have been affected by that schooling. When these last two questions have been asked, it is usually in connection with elite universities or secondary schools.[1]

---

[1] A note on wording: "mobility" should be understood to refer to intergenerational mobility in economic position as measured by occupation and assessed property value.

An extensive sociological literature analyzes both the social determinants of education and the role of education in occupational mobility and in income inequality. Such studies have generally been limited to an analysis of the period in which the research was conducted. Therefore, most of these studies deal with the post-World War II era, and virtually none with the period before World War I.[2]

On the other hand, during the past decade the historical study of social mobility, of which Stephan Thernstrom's work (1964, 1973) is the leading example, has emerged as a major concern of American social historians. This literature is a part of, and an early stimulus to greater interest in, the historical study of American social structure generally.[3] The quantitative information, on which these studies rest heavily, is gathered from sources which list critical socio-economic information about individuals: the United States Census manuscript schedules and comparable manuscript schedules from various state censuses, tax assessment records, city directories listing adult men, marriage licenses, and so forth. In mobility studies, the research also involves tracing a sample of individuals from one source to a second one which provides information about a later point in time.

---

[1] Talbot (1971) is a useful review of recent trends in the history of education generally; Tyack (1975) sums up much of the work on urban education in the period under consideration here, and Lazerson (1973) reviews the "social control" orientation of much of the recent American work in the field.

[2] Jencks (1972) summarizes much of the recent research in these areas. His approach leans heavily on the path modeling of Duncan (Blau & Duncan, 1967; Duncan, Featherman & Duncan 1972) and should be supplemented by Lipset and Bendix (1959), which reviews earlier, less mathematically sophisticated work, by Berg (1970), and by Boudon (1974).

[3] For a convenient collection of examples of the historical study of social structure, see Thernstrom and Sennett (1969). Thernstrom (1973) includes a survey of historical studies of occupational mobility.

These studies have, however, only begun to analyze the relationship between background variables and school attainment, and they have never dealt with the impact of education on mobility.[1]

Rationale. The project proposed here may be seen as an attempt to study these issues by extending the questions of the sociological literature, in appropriately modified form, back in time over a considerably longer period than is now possible. Similarly, it may be seen as an attempt to extend the historical study of social structure to include the examination of education as a variable.

The period 1880-1925 encompasses a substantial increase in the amount of schooling the average American city dweller received. The most obvious indication of this change is the proportion of the population entering secondary school and the consequent transformation of that institution. Within these years it will be possible to examine the role of schooling before and at a number of points during the period of transformation.

---

[1] The sketchy evidence in early twentieth century studies on the social origins of students is briefly described by Krugg (1964). See also the study of G.S. Counts (1922). These studies do not deal with the evolution of the social composition of the schools over time. Troen (1973, 1975: 116-132), Kaestle and Vinovskis (1975) and Katz (1972b) have studied census data on social factors relating to school attendance by age. The first briefly treats the period ending in 1880, the two latter study the mid-nineteenth century. Jensen (1973) is currently studying the impact of length of schooling on career in a rural midwestern state, relying primarily on the 1925 Iowa State Census which asked respondents' length of schooling.

Significance. Throughout American history, education has been regarded as a means of promoting social equality by promoting social mobility and its efficacy in accomplishing this end has been scrutinized in contemporary America. But appallingly little is known about the distribution of education or its role in the stratification process, especially among the poor, in the American past.[1]

This study examines 1) the changing distribution of schooling among various social groups; 2) the extent to which schools helped people to get ahead, particularly those at the bottom of the heap; 3) the extent to which different ethnic groups differed in the use they made of the schools in getting ahead; 4) educational patterns of women of different social backgrounds and the role education played in the social mobility of women; and 5) the extent to which educational credentials required for various jobs changed as the amount of schooling common in the city changed. Results of the research will be carefully compared to results of similar studies in present-day American society.

The distribution of schooling was very different in the 1880-1925 period than in the post World War II era. There is thus reason to think that the role of education in the process of stratification may have also been quite different than it is today. This study substantially extends the time frame in which the evolution of the distribution of schooling and its role in stratification can be studied. Moreover, the period of great secondary school expansion, on which the study focuses, is of particular interest because it provides a basis for comparison with the not dissimilar expansion of higher education in the post World War II era.[2]

---

[1] On the emphasis placed on education as a means of social mobility, see, among innumerable others, Cremin (1961) and Katz (1971). On the contemporary scene, see the references cited in note 2, page 2. On perceptions of the benefits of college for business careers, see Wylie (1954).

[2] See Trow (1961) on this comparison.

SPECIFIC AIMS: QUESTIONS TO BE ADDRESSED

1. General

The issues of the social background of school attainment and the impact of education on mobility will be explored through several subsidiary questions.

Social Background and Educational Attainment. Length of schooling, by social background, will be examined for several points in time extending over half a century.

In general, of course, the upper reaches of society received substantially more education than those below. But how many exceptions were there, extending how far down the social hierarchy?

What was the relative share taken by various social groups in the expansion of the school system during this period? And to what extent were advances at the bottom of the social hierarchy mitigated by tracking?

In order to understand more about life within the schools, it will be instructive to look not only at the distribution of schooling within an age cohort, but also at the related issue of the social composition of educational institutions. How much social class integration, for example, typified the high schools at various times?

Strength of the Relationship between Length of Schooling and Mobility, 1880-1925. When the average length of education was quite brief, one might expect to find a relatively weak relationship between the two, particularly with regard to mobility at the lowest levels of society. However, it may be the case that the fairly small number of children of manual laboring fathers who graduated from grammar school account for a relatively high proportion of those entering the low white collar occupations. In any event, as the amount of education typical in the society at large increased, did the level of schooling of the upwardly mobile increase even more rapidly? To put the question more boldly, did the proportion of men who got ahead without a relatively long education decrease?

Academic Performance and Mobility. Was academic performance, as opposed to length of schooling, related to mobility? Students' average grades and grades in individual subjects (English, mathematics, etc.) are available.

Academic Program and Mobility. Did the particular high school, or the high school program make much difference in the chances for mobility?

Industrial Education and Mobility between Manual Labor Strata. What was the role of industrial education in occupational mobility? Did children of unskilled workers use it as a means to a skilled job (if so, were they more successful, on the average, than others without the education?) or were most of the students in these programs from a skilled labor background? Also, does it appear that to have attained a given rank on the traditional academic ladder (particularly to have finished grammar school) had any impact on mobility within the manual hierarchy?

Later Career Experience. If the better-educated were, at least in some circumstances, more likely to be upwardly mobile, was the difference primarily a result of the early years in the job market? If so, what were the later career experiences? Was there, for example, much chance for the less well-educated to "catch up" in later years? And, among those from manual laboring families whose early occupational experience had made them upwardly mobile into the low white collar stratum, were the better-educated more or less subject to the phenomenon of skidding back to a lower class?[1]

Patterns of Upward Mobility. The particular patterns of upward mobility as well as overall rates will be compared to see whether particular occupational areas were more common among, or limited to, the better educated. The proportion such patterns of mobility comprise among the entire group of the upwardly mobile will be examined.

Educational Credentials as Job Requirements. Was there, in this period of increasing average length of schooling, an upgrading of educational requirements for jobs (Jencks, 1972: pp. 180-185; Lasch, 1973)? A beginning can be made in analyzing their interaction. The levels of education of people entering a range of occupations at intervals during a forty-year period will be available for examination (see also below, page 14).

---

[1] Analysis of later career attainments will help shed light on another issue as well. Assuming that at least for certain social groups, mobility is related to education, the question of causality must be faced. Were the skills necessary to do the job obtained in school; or was an educational credential a test employers imposed to sort their applicants, but which bore little relation to educational skills needed to do the jobs; or, finally, were people who succeeded at school the kind who were likely to succeed in some other situations as well (because of their intelligence, docility or some other qualities)? The available data will obviously not permit definitive answers. But the credentializing effect of schooling can at least be probed if one assumes that its effect is primarily on one's early career and that on-the-job performance is increasingly important thereafter. It is worth making this assumption at least long enough to examine whether the relation between education and career attainment becomes substantially weaker as men grew older (cf Jencks, 1972: 182; Blau and Duncan, 1967: 180).

The Role of Other Family Background Variables. It will be possible to assess the strength of the relationship between many family characteristics on the one hand and educational and occupational attainment on the other (e.g. the impact of nuclear vs. extended family situations, the age of the parents, the absence of a parent, the number of siblings, and the sibling position of the sample members). The extent of unemployment of parents and the class of worker (whether self-employed) will add further dimensions of economic status. Information on parental literacy will be examined as a crude indicator of parental educational attainment. Finally, information on the type of neighborhood will be available.

Total Effects of Family Background. Family background influences educational and career attainment in ways that cannot be understood by focusing on any single one, or even any combination, of the variables on which information is available. But while all the specific ways in which family is related to attainment cannot be stated, the full extent of that relationship can be explored. Sufficient numbers of pairs of brothers, very similar in age, have been selected in the samples to allow for comparisons of how much more alike the attainment of brothers is than the attainment of individuals alike on specific background variables (e.g. on social class and educational attainment) [1]

Value of Education During the Depression. Those who went to school during much of the period under consideration here passed through the Great Depression of 1929 at some point in their careers. Among those already advanced in their careers, did the better-educated suffer any less skidding than others, protected as they were by better skills or at least by credentials? Or, alternatively, were the jobs in which they found themselves especially vulnerable? And among the those just entering the job market, was the value of a better than average education more or less important for getting ahead than in the years before the Depression?

---

[1] Jencks (1972: 320-50)discusses one method for usch comparions.

2. The Ethnic Dimension

Differences among Immigrant Groups in Educational Attainment. To speak of the children of American city dwellers in the period 1880-1925 is as often as not so speak of the children of immigrants. Several immigrant groups, most notably East European Jews and Italians,[1] appear to have differed substantially in their responses to the schools. A good deal has been written about the cultural traditions of those groups in an effort to explain these differences. And a good deal has been written by way of response that the differences are largely reducible to class differences.[2] The school careers of Jews and Italians of comparable class origins will be examined.

Education and Mobility: Native Whites and Immigrants. Was the relationship of education and mobility different among the immigrants' children than among those of native whites? Did they perhaps rely more on educational attainment to get ahead?

Education and Mobility: Jewish and Italian Immigrant Patterns. Jews were considerably more upwardly mobile than Italians of comparable class background (Thernstrom, 1973: 140-142). To what extent is this difference reducible to different educational attainments (if any)? Also did upwardly mobile Italians and Jews enter the same sorts of occupations to become upwardly mobile? If not, were distinct patterns of upward mobility related to any differences in educational experience? To what extent were the differences in Jewish and Italian performance created by a relatively small number of Jews continuing on to college and entering high white collar professions?

---

[1] For the best summary of the literature, see Olneck (1973).

[2] See the discussion and references in Olneck (1973), Cohen (1970), Greer (1972), and Steinberg (1974).

The Parochial Schools. The Catholic immigrants and their descendents populated the parochial schools. Here again appallingly little is known about the social composition of these schools in this period of rapidly increasing Catholic immigrant population.[1] And nothing is known about the role of these schools, in the late nineteenth and early twentieth centuries, in the mobility of the Catholic immigrants. Was there any significant difference, for example, in the average length of schooling (when class and ethnicity are accounted for) among the Catholics in public and parochial schools? And was there any difference in rates of upward social mobility between individuals (of similar social backgrounds and length of schooling) who attended parochial schools and those who attended public schools?

Education and Mobility: The Black Pattern. Timothy Smith (1972) has studied 1910 census data on the proportion of children attending school by age and race. He has called attention to the relatively long schooling of northern Blacks in big cities. It will be possible to analyze this phenomenon in greater detail and over a longer period than the published census tabulations permit.

---

[1] The U.S. Immigration Commission survey (1911) does provide information on ethnic, if not class, background of parochial students in 1907.

Smith suggests that because of racial prejudice, education may have helped Blacks less than Whites. It may well be that Blacks did not reach the same occupational levels as Whites with comparable class and educational backgrounds. But for the same reason of prejudice it may be the case that among upwardly mobile Blacks the proportion of uneducated was lower than among Whites--that to get ahead at all, Blacks needed education more than Whites.

3. Women

Sex Differences in Educational Attainment. Throughout the late nineteenth and early twentieth centuries girls received consistently longer schooling than boys. At every grade level --below college-- girls dropped out less frequently. In the public high schools between 1870 and 1920 girls constituted 56-58% of the enrollment and 60-65% of the graduates.[1] Contemporary educators were well aware of these patterns and some viewed them with alarm. (Krugg, 1964:171-2).

The choice of parents (and children) to keep girls rather than boys in school for a few extra years is a curious one in a society which 1) is supposed to have seen education as a vehicle to mobility and 2) presumably placed a higher value on the occupational preparation of boys. The choice may reflect the workings of the job market for young people: relatively good jobs may have been available for boys but not for girls. The choice may also have reflected cultural values: a fear for the good name and good character of girls who went to work too early. By analyzing the social background of educational attainment among women as well as among men, this study will help clarify which social groups made this choice.

---

[1] Sources for the educational attainment of the American population are reports by schools to the Office of Education, census figures on the proportion attending school at a given age, and census figures (from 1940 only) on educational attainment by age cohort. All indicate longer schooling for girls. The figures quoted here are from the USOE Biennial Report of Education (1930). On the difficulties of reconciling the three kinds of sources, or using any one, see Folger and Nam (1967), B. Duncan (1968) and Jencks and Riesman (1968) as well as references cited there. All three sources, however, indicate longer schooling for girls. Krugg (1964), Tyack (1975), Troen (1975) and Davies (1974) have called attention to the longer schooling of girls in the period 1880-1925. So too has B. Duncan, in connection with the earliest of the cohorts which she analyzes (aged 17 in 1917-21). See especially her tables of continuation ratios and the proportions completing 12th grade.

Implications of Enrollment Patterns for Careers. Whatever the causes for the pattern of the predominance of girls in the high schools, if education paid off in the job market, that pattern may have had ramifications for both sexes. Those social groups deciding (or forced by economic conditions) to make the differential investment in the education of girls may have "invested unwisely" if the women left the labor market after a few years to become housewives while the men remained in the labor market with poorer educational credentials than they might have had.

On the other hand, it should be recalled that in the period under consideration, clerical occupations constituted the most rapidly expanding sector of the labor force, and that the proportion of women in these occupations rose from a tiny minority to an overwhelming majority (Hooks, 1947; Davies, 1974). Similarly, other white collar occupations were expanding and were becoming clearly identified as women's occupations, such as teaching, nursing and library work.

Women with an education may well have found their ways into these jobs, making important financial contributions to their families. Moreover, by virtue of the work environment which they entered, they may have become more likely to marry white-collar men. Because samples of women will be traced through a unique set of marriage records (discussed below) it will be possible to assess the relative importance of occupational mobility, and of mobility by virtue of marriage, the possible facilitation of the latter by the former, and the relation of education to each.

Thus, the pattern of differential investment in the education of girls among some social groups may have prompted upward mobility among women rather than among men. To restate the same point from another point of view: education's greatest impact on intergenerational mobility into white collar,and especially low white collar, occupations may have been among women.

Women: the Ethnic Dimension. It seems reasonable to expect that the position of women in different ethnic subcultures led to different ethnic patterns of female educational attainment and mobility. It will be important to compare, for example, not only the experiences of Italian men and women, and Black men and women, but also the experiences of Black and Italian women.[1]

---

[1] The analysis of women's experiences must be regarded as exploratory and the questions to be addressed as more tentative than in the case of men. The sociological literature on the mobility of women, despite increased recent attention (a useful bibliography appears in Featherman and Hauser, 1975) remains minute by comparison to the study of male mobility patterns. And historical studies of women's mobility patterns have not been published. Indeed, this project would be breaking new ground by simply reporting on intergenerational mobility patterns of women even if it ignored the educational issues altogether.

References Cited in the Introduction

Berg, I.
1970 Education and Jobs: The Great Training Robbery. New York: Praeger.

Blau, P. and O.D. Duncan.
1967 The American Occupational Structure. New York: Wiley.

Boudon, R.
1974 Education, Opportunity and Social Inequality: Changing Prospects in Western Society. New York: Wiley.

Cohen, D.K.
1970 "Immigrants and the Schools". Review of Educational Research 40: 13-28.

Counts, G.S.
1922 The Selective Character of American Secondary Education. Chicago: University of Chicago.

Cremin, L.A.
1961 The Transformation of the School: Progressivism in American Education 1876-1957. New York: Alfred A. Knopf.

Davies, M.
1974 "Women's Place is at the Typewriter: the Feminization of the Clerical Labor Force". Radical America 8 no. 4:1-28.

Duncan, B.
1968 "Trends in Output and Distribution of Schooling". In E.B. Sheldon and Wilbert E. Moore, eds. Indicators of Social Change. New York: Russell Sage Foundation.

Duncan, O.D.
1961 "A Socioeconomic Index for All Occupations". In A.J. Reiss, ed. Occupations and Social Status. New York: Free Press.

Duncan, O.D., D.L.Featherman, and B. Duncan.
1972 Socioeconomic Background and Achievement. New York: Seminar Press.

Edwards, A.
1943 Comparative Occupational Statistics for the United States, 1870-1940. Washington, D.C.: U.S. Government Printing Office.

Featherman, D.L. and R.M. Hauser.
1975 Sexual Inequalities and Socioeconomic Achievement in the U.S. 1962-1973. Institute for Research on Poverty Discussion Paper #275-75. Madison: University of Wisconsin.

Folger, J.K. and C.B. Nam.
1967 Education of the American Population. Washington, D.C.: U.S. Government Printing Office.

Greer, C.
1972 The Great School Legend: A Revisionist Interpretation of American Public Education. New York: Basic Books.

Hershberg, T. et.al.
1974 "Occupation and Ethnicity in Five Nineteenth-Century Cities: A Collaborative Inquiry". Historical Methods Newsletter 7:174-216.

Hodge, R.W. et.al.
1964 "Occupational Prestige in the United States 1925-63". American Journal of Sociology 70:286-302.

Hooks, J.M.
1947 Women's Occupations Through Seven Decades. Women's Bureau Bulletin #218. Washington, D.C.: Government Printing Office.

Jencks, C. and D.Riesman.
1968 The Academic Revolution. New York: Doubleday.

Jencks, C. et.al.
1972 Inequality: A Reassessment of the Effects of Family and Schooling in America. New York: Basic Books.

Jensen, R.
1974 "Education and Social Structure: An Historical Study of Iowa". In U.S. Department of Health, Education, and Welfare, National Institute of Education, Career Education Program, Program Plan for Fiscal Year 1975. Washington, D.C.: U.S. Government Printing Office.

Kaestle, C. and M. Vinovskis.
1974 "Quantification, Urbanization, and the History of Education: An Analysis of the Determinants of School Attendance in New York State in 1845". Historical Methods Newsletter 8 no. 1:1-9.

Katz, M.B.
1969 "Conspicuous Consumption". In The Hamilton Project: Interim Report No. 1. The Ontario Institute for Studies in Education (duplicated).

Katz, M.B.
1971 Class Bureaucracy and Schools: The Illusion of Educational Change in America. New York: Praeger.

Katz, M.B.
1972a "Occupational Classification in History". Journal of Interdisciplinary History 3:63-88.

Katz, M.B.
1972b "Who Went to School?" History of Education Quarterly 12:432-454.

Katz, M.B., and J. Tiller.
1972 "Record Linkage for Everyman: A Semi-Automated Process". Historical Methods Newsletter 5, No. 4:10-14.

Knights, P.R.
1969 "City Directories as Aides to AnteBellum Urban Studies". Historical Methods Newsletter 2, 4:1-10.

Krugg, E.
1964 The Shaping of the American High School, 1880-1920. Madison: University of Wisconsin Press.

Lasch, C.
1973 "Inequality and Education". New York Review of Books: 20, no. 8:19-28.

Lazerson, M.
1973 "Essay Review on American Educational History". Harvard Educational Review 43:269-283.

Lipset, S.M. and R. Bendix.
1959 Social Mobility in Industrial Society. Berkeley: University of California Press.

Olneck, M.
1973 Immigrant School Achievement 1900-1935. Harvard Graduate School of Education Qualifying Paper.

Rhode Island Office of the Commissioner of Labor
1907 The 1905 Rhode Island Census. In Twentieth Report of the Commissioner of Labor. Providence: State Printing Office.

Shorter, E.
1971 The Historian and the Computer: A Practical Guide. Englewood Cliffs: Prentice Hall.

Smith, T.
1972 "Native Blacks and Foreign Whites: Varying Responses to Educational Opportunity in America 1880-1950". Perspectives in American History VI:309-339.

Steinberg, S.
1974 The Academic Melting Pot. Catholics and Jews in American Higher Education. New York: McGraw-Hill.

Stephenson, C.
1974 "Tracing Those who Left: Mobility Studies and the Soundex Indexes to the U.S. Census". Journal of Urban History I, no. 1:73-84.

Talbot, J.E.
1971 "The History of Education". Daedalus 100, no. 1:133-152.

Thernstrom, S.
1964 Poverty and Progress: Social Mobility in a Nineteenth Century City. Cambridge: Harvard University Press.

Thernstrom, S.
1973 The Other Bostonians: Poverty and Progress in the American Metropolis, 1880-1970. Cambridge: Harvard University Press.

Thernstrom, S. and R. Sennett (eds.)
1969 Nineteenth Century Cities: Essays in the New Urban History. New Haven: Yale University Press.

Troen, S.
1973 "Popular Education in Nineteenth Century St. Louis". History of Education Quarterly 13:23-40.

Troen, S.
1975 The Public and the Schools: Shaping the St. Louis System, 1838-1920. Columbia, Missouri: University of Missouri Press.

Tyack, D.
1974 The One Best System: A History of American Urban Education. Cambridge: Harvard Univeristy Press.

U.S. Bureau of the Census.
1950 1950 Census of Population: Alphabetical Index of Occupations and Industries. Washington: U.S. Government Printing Office.

U.S. Immigration Commission.
1911 The Children of Immigrants in the Schools. Reports. Vol 29-33. Washington, D.C.: U.S. Government Printing Office.

Winchester, I.
1970 "The Linkage of Historical Records by Man and Computer". Journal of Interdisciplinary History 1:107-24.

Wrigley, E.A. (ed.)
1973 Identifying People in the Past. London: E. Arnold.

Wylie, I.
1954 The Self-Made Man in America. New Brunswick: Rutgers University Press.

THE RESEARCH DESIGN AND THE DATA COLLECTION

I. AN OVERVIEW OF THE RESEARCH DESIGN

Between the years 1880 and 1925 the length of schooling received by the average American city dweller increased considerably. The most obvious indication of this change is the proportion of the population entering secondary school, and the consequent transformation of that institution. The research design attempts to capture information about individuals who passed through the school system at different moments in time during this period.

Providence was selected for study as reasonably typical of eastern American urban centers of the period. It was among the 30 largest cities in the nation during the period under study. Its rapidly growing population (104,000 in 1880, 238,000 in 1920) included significant numbers of Irish, Italians, Russian Jews, blacks and other immigrant groups. Its economy based on textiles, jewelry and machine manufacture, also included the area's commercial and financial center as well as a state bureaucracy (albeit of the smallest state). In size, composition and economic structure, it is more a major American city than a mill town. Its schools - including a rapidly developing high school system, primary and grammar schools, a Catholic school system which enrolled about a fifth of the city's pupils, and a small group of prestigious private schools - as well as the educational issues of concern in the city - such as industrial education, high school growth, the Gary Plan, Americanization, IQ tests, the size of the school committee - also make the city representative. The quality of its records, which will be discussed below, is excellent.

Samples of individuals were chosen from four specific years during the period 1880-1925. Manuscript schedules are available from censuses taken by the federal government in 1880 and 1900 and by the State of Rhode Island in 1915 and 1925. These schedules were the source from which samples of young

---

This research was supported by grants from the National Institute of Education and the Center for the Study of Metropolitan Problems of the National Institute of Mental Health and by the Joint Center for Urban Studies of MIT and Harvard University.

people were drawn. These sample members were then traced to school records to determine the length and the nature of their schooling. Finally, all the males were then traced forward in time to city directories (which indicate their occupation) and to tax books (which indicate the value of their property, the only other available measure of economic attainment). They could not be traced all over the country; they were traced to the city directories of Providence, and to the directories of those nearby towns to which they would have been most likely to move had they remained in the area.

Samples of women were also selected from the censuses in order to shed light on patterns of schooling and of economic attainment among women. But these samples were drawn only from the records of 1880 and 1915 on the assumption that information from two points in time would be sufficient to highlight distinct female patterns.[1] The women were also traced to the school records and forward in time. But the latter trace could not be made to the city directories since these volumes regularly omitted large proportions of women and since even those included might not be found under their maiden names. Hence the women were traced instead to marriage records, which also provide information on whom they married.

Traces to other sources, and the collection of other samples, were all simply variations on this basic design: an age cohort selected from census records, traced through school records and traced forward in time to city directories and tax books or to marriage records. The first variation was to trace all sample members' guardians to the city directories and tax books of the census year. Additional information was thus obtained about their occupation, property holdings and also about the spelling of their names (critical for tracing across time). The second, and more extensive, variation involved collecting samples to investigate special populations more closely.

There are several such samples:

1) High School Samples. Most adolescents did not attend high school at all during most of the period under consideration. In order to investigate the high school population in detail, a separate sample of high school entrants in the census years was selected from high school records. It was traced back to the census records, and then to other sources in the same way as the census sample.

2) Ethnic Oversamples. There will be too few blacks in the census samples to study in any detail. Since their relation to educational institutions is a matter of considerable interest, separate samples of blacks in the same age ranges as the random sample were chosen.

I was also interested in studying immigrant groups in detail. Considerable differences in their responses to the school system have been widely claimed and heatedly denied. I therefore selected large oversamples of Jews and Italians from the census manuscripts of 1915. Finally, I selected an oversample of Irish children from this census, both in order to be able to compare the later arrivals of this important group to earlier Irish immigrants and to compare them to the experiences of other groups in 1915.

3) The 1880 Sample of the Unschooled. School records for the 1880 period do not exist for those who did not complete grammar school - which is to say for the majority of the population. In order to subdivide this majority in some way according to their educational background, I exploited the census question "attended school in the last year?". Some 80% of 12-year-olds and 90% of 11-year-olds were reported to have attended. I therefore collected a sample of those who had not attended. This group may be thought of as constituting the least well educated portion of all who did not complete grammar

school, and can instructively be compared to the rest.

4) The Marriage License Sample of White Collar Brides. Because education was hypothesized to be particularly valuable to the mobility of women, for whom it may have served as an entry to white collar clerical and sales jobs, a special sample of white collar brides was selected from the marriage records. These women were traced backward in time to the school records.

* * *

Table I-1 provides a summary of these samples and the number of individuals in each. Figure I-1 is a summary of the sources from which sample members were selected and to which they were traced. Each subsequent section of the paper describes in detail the nature of the sources and the rationale for the specific way in which they were used. Generally, the most technical issues, minor exceptions to statements in the text, and anticipations of some questions appear in the notes rather than in the body of the paper. But the distinction is not, of course, a precise one. Finally, it should be noted that this paper describes the quantitative evidence gathered especially for this study, evidence on which the study will rest heavily. The paper does not deal with the use that will be made of other, more traditional sources.

Table I.1 The Size of Each Sample

SELECTED FROM THE CENSUS PERIOD	SAMPLE	Number included in sample of Men	Number included in sample of Women
1880	11-16 year olds in Providence	1223	898
	11-16 year old Blacks	125*	142*
	11-12 year olds not attending school	249*	207*
1900	13-16 year olds in Providence	914	
	13-16 year old Blacks	126*	
1915	12-15 year olds in Providence	915	903
	12-15 year old Blacks	155*	158*
	12-15 year olds whose fathers were born in Russia	388*	324
	12-15 year olds whose fathers were born in Italy	581	291
	12-15 year olds whose fathers were born in Ireland	446	215
1925	12-15 year olds in Providence	1090	
	12-15 year old Blacks	165*	
SELECTED FROM THE PUBLIC HIGH SCHOOL RECORDS			
	Individuals entering High School in		
	the 1880 period	216**	112**
	the 1900 period	581**	367**
	the 1915 period	349**	416**
SELECTED FROM THE MARRIAGE RECORDS			
	Brides with white collar occupations who were born in the same years as the 1915 Census sample		492

* Includes entire population in age range - not a sample.
** Includes entire classes in selected years - not a sample.

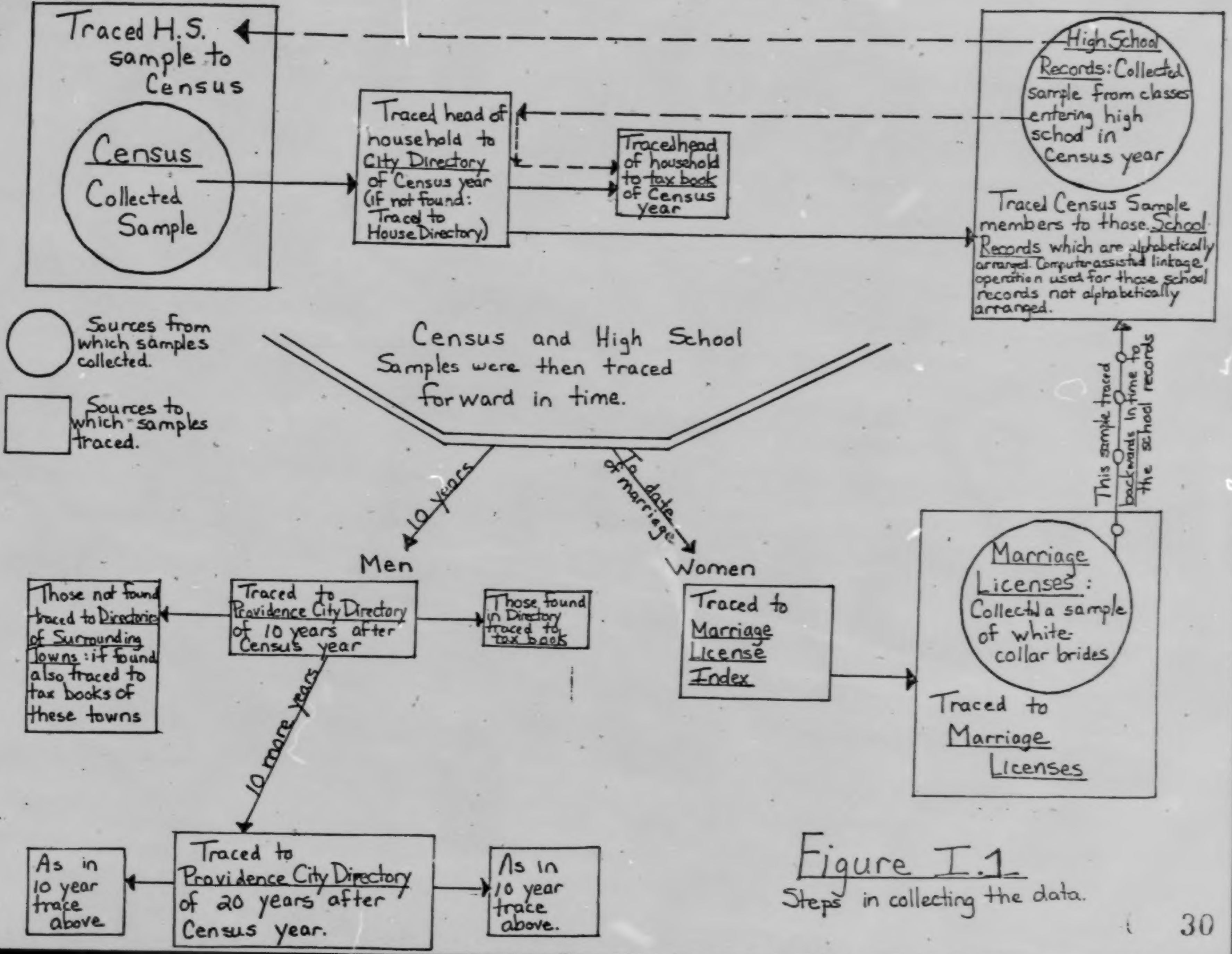
Traced H.S. sample to Census
Census
Collected Sample
Traced head of household to City Directory of Census year (if not found: Traced to House Directory)
Traced head of household to tax book of Census year
High School Records: Collected sample from classes entering high school in Census year
Traced Census Sample members to those School Records which are alphabetically arranged. Computer assisted linkage operation used for those school records not alphabetically arranged.
Sources from which samples collected.
Sources to which samples traced.
Census and High School Samples were then traced forward in time.
10 years
To date of marriage
This sample traced backwards in time to the school records
Men
Women
Those not found traced to Directories of Surrounding Towns: if found also traced to tax books of these towns
Traced to Providence City Directory of 10 years after Census year
Those found in Directory traced to tax books
Traced to Marriage License Index
Marriage Licenses: Collected a sample of white-collar brides
Traced to Marriage Licenses
10 more years
As in 10 year trace above
Traced to Providence City Directory of 20 years after Census year.
As in 10 year trace above.

Figure I.1

Steps in collecting the data.

II. SELECTING THE SAMPLES

A. THE CENSUS SAMPLES

Census Underenumeration. The Census is the obvious - indeed the only - source for the selection of a teenage cohort of the population. The Manuscript schedules prepared by the Census takers have been used and described by numerous scholars. On the other hand, evidence on the completeness of early Censuses is virtually nonexistent. Peter Knights has estimated that perhaps 8% of the population was omitted in early nineteenth century Censuses of Boston.[1] But that estimate exhausts the research on the subject. Besides assuming that the lowest social classes were most likely to be missed by the Census taker, and attempting to guage what effect this might have had on particular results, there is in fact little one can do to deal with the problem of potential census underenumeration.

Sample Selection. In order to obtain a representative sample of adolescents, I chose individuals from the entire city rather than only from selected neighborhoods. For the same reason I made no attempt to exclude individuals with common names. It proved extremely difficult to trace such individuals across time and even to some school records (these issues are discussed below). Nevertheless it seemed preferable to avoid the biases that might be involved in limiting samples to those with uncommon names.[2]

The age range of the sample members was chosen in order to maximize the probability of successfully tracing them to records of grammar school graduation or high school entrance. Therefore, the three most common ages of grammar school graduation were chosen as well as the age one year younger than these to increase the size of the group sampled. In the 1880 and 1900 periods these ages turned out to be 13-16 and in the 1915 and 1925 periods 12-15. The change is due to the fact that the grammar school course was shortened by one year in 1904.

However, for the 1880 period the age range was extended to include individuals 11-16 years old. This was done in order to be able to make comparisons between a randomly selected group of 11 and 12 year olds and the special sample of the unschooled. The latter, it will be recalled, includes the small minority of 11 and 12 year olds who were not in school in 1880. Extending the age range by two years was also advantageous because it increased sample size by half. This increase is particularly valuable in the 1880 period because a higher proportion of the sample members in that year had common names and consequently may have to be excluded from the sample that can be traced across time.

The method of selecting sample members should ideally allow each relevant individual an equal chance of falling into the sample (in this case each individual in the age range). Such a sample has the best chance of adequately representing the population from which it is drawn. Moreover, statistical procedures generally assume such a sample. Individuals selected with the aid of a random number table (a table in which every digit has an equal chance of occurring in every position) would constitute a true random sample. However, only about 2% of the population fell into the 4-year age range in which I was interested. Therefore, a pure random sample (based on, say, combinations of randomly chosen Census enumeration districts, page numbers, and lines) would have involved rejecting some 98% of the lines selected. This seemed to me an unnecessarily inefficient procedure. A common alternative to a pure rendom sample is a systematic sample, in which every Nth individual is chosen. But for my purposes such a procedure had exactly the same drawback as a true random sample. I chose instead to sample every Nth page of the Census and to include every individual in the age range on sampled pages. Since there were usually 50 individuals listed on each page, there tended to be very few sample members chosen from any given page.[3] Indeed, on average, one was chosen from each page. Strictly speaking, the samples should probably be defined as cluster samples (each page constituting a cluster chosen by systematic sampling) but with the number of clusters so high and the number of sample members selected from each cluster so low, it makes more sense, I think, to consider the sample as systematic.[4] ¶ When two (or more) siblings in a family fell into the appropriate age range, both were included in the sample, It will thus be possible to compare the experiences of siblings to some extent, and, specifically, to investigate how much more similar they are than individuals who share certain identifiable background traits such as social class and ethnicity.[5]

The Ethnic Oversamples. In all four periods, all Black boys in the age range were included in ethnic oversamples. In 1880 and 1915, the years for which random samples of girls were collected, information on all Black girls in the age range was also collected.

In 1915 oversamples were also selected from the Irish, Italian and Jewish populations. The number of girls collected was somewhat smaller than the number of boys / in order to conserve time and money, and in the hopes that, coupled with information on the boys, the ethnic dimension could still be adequately explored. I also assumed that there was no point in selecting oversamples of those groups in other periods since the basic comparisons could be made from one point in time and because the one group whose position changes greatly - the Irish - are adequately represented in the 1880 and 1900 random samples.[6]

The criteria for inclusion in the Black oversamples depended, of course, on the race of the individual. Those classified "mulatto" were also included, but given a separate code. Inclusion in the Irish, Italian and Jewish oversample was based on the place of birth of the father. The sample therefore includes first and second generation group members in the correct proportions.

All whose fathers were born in Russia were included in the Jewish oversample. The Census does not include a question on religion, and did not distinguish among ethnic groups having the same place of birth (for example between Russian, and Russian Hebrews). It is therefore necessary to rely on place of birth as a proxy for East European Jews. Many studies have used the designation "Russian" in this way because the overwhelming majority of immigrants from Russia were Jews. That this was so in Providence is clear from evidence from the 1905 Rhode Island state Census. This Census did in fact ask about religion. And 95% of those Providence residents born in Russia were Jews. Moreover, the non-Jewish Russians were clustered in one ward. While no figures are available for 1915, there is no reason to assume the proportion had changed in the intervening decade. And if a study of the neighborhoods involved suggests no major changes between 1905 and 1915, it may also be possible to refine the sample by excluding clear-cut cases of non-Jewish Russians, with the help of the 1905 Census descriptive data and a close look at family names and given names.[7]

The exclusion of Jews from other countries from the oversample is unimportant, since my purpose is to compare different immigrant patterns of behavior, not to draw conclusions about all Jewish immigrants. Russian Jews are probably representative of all Jews from Eastern Europe, but even if they were not, the comparison to the Italians would be satisfactory.[8]

To avoid a large duplication of effort, the pages selected for the ethnic oversamples included all pages selected for the random sample. The research assistants then skipped these pages in collecting the oversample. And the computer was later programmed to select from the 1915 random sample all Irish, Italian, Jewish and Black sample members, to duplicate these cases and to include the duplicated cases in the oversamples. Thus, for example, the oversample of 1915 Italian boys was selected from half the pages of the census, less the pages - one in eight - that had fallen into the random sample. Then the computer duplicated all those cases in the random sample whose male guardian had been born in Italy.

Sample Representativeness. The representativeness of the census samples is rather difficult to measure because they are samples of adolescents. There are few tables in the published census data which report figures on adolescents. Nor can their guardians be compared to tabulations since the census does not, of course, report figures on "guardians of adolescents." Even a comparison of all male guardians in a given age range - say 40-55 - to census data on all men 40-55 is problematic. Not only are the groups not strictly comparable, but the number of possible comparisons are relatively few. The federal Census of 1880 and 1900 provided few tabulations by age for adults in cities of 100,000 - 200,000. For the two later periods no tabulations whatsoever are available. Although the Rhode Island State Census Reports for 1885, 1895 and 1905 were very detailed, no such reports were prepared for the 1915 and 1925 Censuses. As a result comparisons with published census data involve ignoring differences in the definition of the groups being compared, or differences in the years from which the information comes, or both. Such comparisons would be useful, if at all, only to establish whether the magnitudes in the samples are roughly about right. I have included two such comparisons because published data make it possible to estimate with some confidence what the population distributions are and hence what the sample distributions should be. In Table II-1 both the number of the sample members and the proportion of them who are foreign-born are compared to estimates of what the values should be. The "explanation" column (and the notes) in the table reveal the estimation required to obtain the expected values from the published figures. The published figures do not pertain to the year from which the sample was selected and in most cases they pertain to a somewhat different age range. Figures were usually available for 5 years before and for 5 years after the year needed. The figures from these two dates, then, were regarded as a range within which was to be found the number for the year sampled. For example, no figure for the number in the age range in the 1915 R.I. Census is available. But figures are available for 1910 and 1920 U.S. Censuses. They form a range within which the 1915 figure should fall. The "estimate" presented is simply the average of the figures for the census occurring immediately before and after the one from which the sample was drawn. In the 1915 case, the estimate is the average of the figures from 1910 and 1920.

Thus the range is really more meaningful than the estimate. There is, after all, no guarantee that in 1915 the city was just at the midpoint in its demographic evolution between the 1910 and 1920 circumstances. Moreover, the greater reliability of the range is particularly relevant in examining the absolute number of cases expected, as opposed to the nativity distribution (the distribution will not change as quickly as the absolute numbers). The actual number in the sample is indeed within the range in every case but the last. And in 1925 the population of the city was greater than in either 1920 or

1930 (see note 3 to table) so that the number in the sampled age range should indeed be outside the range. The distributions by nativity not only fall within the ranges in every case for which a range is available, but they are also impressively close to the figures in the estimate column. In none of the six samples does the actual distribution differ significantly from the estimated one.

But Table II-1 also demonstrates the difficulty of obtaining an estimate from the published reports against which to make even the simplest test of the data. The estimating procedure required for checking other variables would be more cumbersome and probably less accurate. However, another way of assessing representativeness is by comparing the samples of girls and boys in 1880 and 1915. If the distributions of each are similar, it is far less likely that both are unrepresentative than that both are representative.[9]

In Table II-2 the sample members have been compared in terms of a number of characteristics. Since girls and boys might differ in many ways (proportion at school, at work, etc.) the only direct comparisons between male and female sample members are on ethnicity and on the proportion having a sibling of the same sex in the sampled age range. On the other hand, their guardians or their households have been compared on six additional characteristics. In none of these comparions - a total of fifteen of them - were the samples found to differ significantly.

The information obtained from the census manuscripts is summed up in the form (appended at the end of this paper).[10] The information on the sample members themselves was recorded almost exactly as it appeared. All the available information on other household members - parents, siblings, other relatives, servants, boarders - might also have been collected. But a large amount of time and effort would have been required for information which included a good deal I would probably never use. I therefore chose instead to record information in full only on the sample member and his

Table II-1 A Comparison of Actual and Expected Distributions of Sample Members, by Nativity

PERIOD	SAMPLE	NATIVITY	Actual Distribution %	Expected Distribution		Explanation
				Estimate %	Range %	
1880	Males	Natives	87.6	88.3	-	From the R.I. State Census of 1885, adjusted for 1880[1]
		Foreign born	12.4	11.7		
		Total	100	100		
		(N)	(1223)	(1336)		
	Females	Natives	85.7	87.9		
		Foreign born	14.3	12.1		
		Total	100	100		
		(N)	(898)	(966)		
1900	Males	Natives	81.2	81.3	82.1-80.7	Range from R.I. censuses of 1895& 1905; Estimate is their average
		Foreign born	18.8	18.7	17.9-19.3	
		Total	100	100	100	
		(N)	(914)	(951)	(800-1102)	
1915	Males	Natives	84.6	86.5	81.9-90.8	Range from U.S. Censuses of 1910 and 1920; estimate is their average[2]
		Foreign born	15.4	13.5	18.1-9.2	
		Total	100	100	100	
		(N)	(915)	(918)	(884-952)	
	Females	Natives	84.4	87.0	82.6-91.0	
		Foreign born	15.6	13.0	17.4-9.0	
		Total	100	100	100	
		(N)	(903)	(933)	(883-982)	
1925	Males	Natives	94.8	93.7	90.8-96.4	Range from U.S. Censuses of 1920 and 1930: estimate is their average[2]
		Foreign born	5.2	6.3	9.2-3.6	
		Total	100	100	100	
		(N)	(1090)	(1013)	(952-1073)[3]	

Note: None of the differences between the estimated expected distributions and the actual distributions are significant at the .05 level or at the .1 level. At the .25 level three would be significant (1880 females and both 1915 samples).

Table II-1 (continued)

1. Adjusted: $\frac{\text{total city population, 1880}}{\text{total city population, 1885}}$ x the figures by age and nativity for 1885

2. The U.S. Census reports for 1910-1930 do not provide the numbers of 12-15 year olds. They report on 10-14 year olds. The ratio of 12-14 year olds to 10-14 year olds in each sex and nativity group in the R.I. State Census of 1905 were used to estimate the number of 12-15 year olds in each category in 1910-1930.

3. The actual sample size does not quite fall within the range expected. But both the 1920 and 1930 U.S. Censuses reported smaller populations for Providence than the 1925 R.I. State Census did. The figures are: 1920: 237,595; 1925: 268,400; 1930: 252,981. If the city's population was larger in 1925 than in either 1920 or 1930, the number in the sampled age range no doubt also was larger. No report of the 1925 R.I. State Census was published; however the total for Providence was published in Providence City Registrar, 86th Annual Report Upon Births, Marriages and Deaths in the City of Providence for 1940 (Providence: City Printer, 1940), 128.

Table II-2. Comparisons of the Frequencies Obtained in the Sample of Males and the Sample of Females (1880 and 1915)

	1880		1915	
	Sample of males	Sample of females	Sample of males	Sample of females
A. Ethnicity of Sample Member	%	%	%	%
Black	2.1	2.8	2.5	2.4
white, father born in U.S.A.	38.6	36.8	32.7	34.3
Canada	2.4	2.6	7.2	5.8
Ireland	40.6	40.6	12.6	13.5
England, Scotland, Wales	11.6	13.4	8.0	7.8
Eastern Europe			10.0	11.1
Italy			14.4	15.5
other	4.6	3.4	11.3	8.9
total	100	100	100	100
(N)	(1223)	(898)	(915)	(903)
B. Identity of Head(s) of Family				
Both parents	74.0	74.4	78.6	76.0
Two other relatives	3.1	3.8	3.7	3.4
Father only	4.1	5.1	3.5	3.5
Mother only	5.3	4.8	4.3	5.3
Mother and an adult male relative	7.6	5.8	6.0	5.9
No relative	4.2	3.7	2.1	2.3
other	1.8	2.4	1.9	3.5
total	100	100	100	100
(N)	(1223)	(898)	(915)	(903)
C. Number of sample members per household in Sample[1]				
1	82.1	79.5	87.5	86.0
2 or more	17.9	20.5	12.5	14.0
total	100	100	100	100
(N)	(1037)	(745)	(813)	(792)
D. Number of servants in father's household[2]				
0	89.4	93.0	96.2	96.9
1	7.2	5.2	3.2	2.4
2 or more	3.4	1.7	.6	.6
total	100	100	100	100
(N)	(794)	(574)	(660)	(621)
E. Total number of lodgers in father's household[2]				
0	90.8	90.2	91.4	89.9
1	6.3	5.9	5.9	5.8
2 or more	2.9	3.8	2.7	4.3
total	100	100	100	100
(N)	(794)	(574)	(660)	(621)

Table II-2 (continued)

	1880		1915	
	Sample of males	Sample of females	Sample of males	Sample of females
	%	%	%	%
F. Father's Occupation[2]				
High white collar	5.5	5.1	2.9	4.7
other managers, officials, and proprietore	9.2	10.6	11.5	10.0
other white collar	8.7	6.5	17.6	16.9
skilled manual	29.7	30.1	24.7	22.9
semi-skilled manual	22.4	26.1	24.7	25.6
unskilled manual	22.2	20.6	17.3	19.0
other	2.1	.7	1.4	.8
total	100	100	100	100
(N)	(794)	(574)	(660)	(621)
G. Assessed Value of Father's Property[3]				
0	69.0	66.8	70.0	69.3
$10-$1,000	7.9	9.1	8.1	8.0
$1,001-$2,500	6.9	8.1	7.1	6.9
$2,501-$5,000	6.1	8.1	6.1	8.0
$5000-$15,000	5.1	5.0	7.1	6.9
over $15,000	5.1	3.0	2.0	.8
total	100	100	100	100
(N)	(493)	(397)	(495)	(476)
H. Literacy of Father[2]				
can read and write	87.4	87.3		
other	12.6	12.7		
total	100	100		
(N)	(794)	(574)		

NOTE: None of the x^2 values for these crosstabulations is significant at the 0.5 level. Only 3 (C, D and F in 1880) are significant even at the .25 level.

1. N = number of households in sample.

2. These tables include only cases in which the father was present and head of household. Also each father is counted only once even if two or more of his children were included in the sample.

3. In addition to the restrictions noted in footnote 1, only fathers who could be traced unambiguously to the books were included (see text, Section III B).

parents. In addition the number of older and younger siblings of each sex was recorded. Hence the sample member's birth order position can also be determined. For other relatives I recorded only the number by generational status. I also recorded the number of servants and the number of boarders.

Finally, I added one key variable: the child-guardian relationship. It reflects 1) whether both parents, one, or neither are present, 2) if only the mother is present whether some other adult male relative is present, and 3) the relationship of the child to the head of the household (who is not always a parent). Thus the information on household members, together with the child guardian relationship code, allows for classification of household membership and structure along most of the dimensions used by family historians - size, presence of boarders, presence of servants, whether female headed, whether nuclear or extended, if extended, how and so forth. At the same time *it* locates the sample member himself in this structure. He, not the family, remains the unit of analysis.[11]

B. THE PUBLIC HIGH SCHOOL SAMPLE

The quality of the school records, and the nature of the information that could be gathered on schooling will be discussed in detail below. This section concerns only the selection of the public high school samples and the rationale behind it.

The information on entrants to the Catholic high schools and the other private secondary schools in the city was not sufficiently complete or sufficiently detailed to merit selection of separate samples from these institutions. Hence the samples were drawn only from the public high schools.

In the 1880 period there was a single public high school in the city - the Providence High School. It included Classical, English and Science, and Girls' departments. By the 1900 period, this institution had been replaced by Classical High School, English High School, Hope High School (which served a specific region of the city and included the same curricula as the other two) and Manual Training High School, soon to be renamed Technical High School. These same institutions continued throughout the period studied.

The samples from the public high schools include all students entering ninth grade in the year of the parallel census sample. Thus all those entering high school in 1880, 1900 and 1915 were included. Because the number of boys who entered in 1880 and 1900 was too small for an adequate sample, I also included those entering in the year before and the year after.[12] No sample was selected from the 1925 period since the boys in the census sample attended high school in sufficient numbers to make a special sample of high school students unnecessary.

Finally, the records of the Manual Training High School (Technical High School) have been interfiled alphabetically in an enormous file of records, a quarter of a million in number, to which it is easy to trace a person but from which it is difficult to select a sample from a given year. Consequently no special oversample was selected from this high school. The information on entrants to this school will have to be based on the census sample members. Since the school admitted large numbers of boys in 1900 and especially in 1915, the limitation is not too serious. A fair number were included in the census samples. The same is true of girls in the 1915 period. No 1900 census sample for girls is available but the school only admitted a tiny number of girls in that period so the omission should not be very important.[13]

In order to create a sample of all high school students it will be possible to supplement the public high school sample with the census sample members who entered schools not included in the high school sample (Technical

H.S., Catholic high schools and other private high schools). The cases would be weighted on the basis of published enrollment information and on the basis of the number of entrants to each high school found in the census sample.

III. OBTAINING ADDITIONAL INFORMATION ON FAMILY BACKGROUND.

A. TRACING SAMPLE MEMBERS'GUARDIANS TO THE CITY DIRECTORY AND TAX BOOK

The Directories. The city directory lists individuals in alphabetical order,indicating for each one his address and occupation and often his business address. Like the census manuscripts,they have been used and described by various scholars. I have explored their inclusivity and the problem of accurately matching individuals to directory listings in another paper.

As indicated in the overview, the guardians of sample members were traced to the directory of the census year in order to supplement the information on their occupations. In some cases the census had not provided it, for the others it will be possible to compare two reports of occupation. The trace also provided additional information on the spelling of the sample member's surname which often proved helpful in tracing the sample members themselves. Finally, this trace provided information on the commonness of the guardian's name, which proved essential for the use of the tax books.

For each period except 1880, a biennially published house directory was also available. It ordered entries by street and number rather than in alphabetical order by name. It appears to be far less inclusive than the city directory, perhaps limiting listings to heads of households. But it was of considerable use in locating those individuals whose names were hopelessly misspelled in the census (or incorrectly copied from the census). All guardians not successfully matched to a listing in the city directory were traced to the house directory.[1] Since part of the purpose of tracing guardians to the directory was to obtain information for use in other traces (on the spelling and commonness of their names) it was important to be as cautious as possible in matching guardians to listings. The same address was therefore used as the criterion for a valid match.

The directory information was copied exactly as it appeared: occupation, business address (when it appeared), and home address. Information on addresses - whether from the census or from the directories - was not made machine readable. Residence is surely a dimension of attainment worth pursuing, but I have not found a way to easily incorporate information from street addresses. The information could, of course, be coded at a later date and added to the computer tapes.[2]

A second serious omission is information on whether artisans were self-employed. The city directory lists tradesmen by occupation in the business directory(printed in the same volume as the regular alphabetical directory of residents). But a second search is involved here, and I did not have time to include it. The 1915 Census provides information on the self-employment status of sample members' guardians. It will thus be possible to make some evaluation of how serious the omission is. It may also be worth at least checking on the artisans in the 1880 sample, since self-employed artisans may well have been more numerous in the early period. Once again, of course, the information can be gathered later and, once gathered, incorporated into the existing data with relative ease.

In any event, property tax information may make it possible to distinguish artisans who had shops of their own from those that worked for others. Artisans with a certain amount of property could, for example, be assumed to be petty proprietors, and those with less than that amount not. Even if the categorization involved significant numbers of errors, the more important cases - the artisans who really were significantly better off than other skilled workers - would be identifiable.[3]

The Tax Books. The Providence Tax Book has been published annually since the early 19th century. It includes the name of each individual owning taxed property, the assessed worth of that property and the amount of the tax itself. The property is divided into two categories, real property and personal property. The latter includes all property other than real estate. After 1912 it was subdivided into tangible personal property (such as machinery or furniture) and intangible personal property (such as bonds or savings accounts). The actual assessment for any individual is probably only a very rough estimate of his worth. Certain property escaped assessment through loopholes in the laws. Salaries, of course, were not taxed as such, and savings accounts may also have generally not been taxed. Also different assessors may have evaluated property quite differently. Finally, they generally appeared to assess at only a fraction of the value of the property.

Still, as a rough approximation of the wealth of an individual, the assessed property values may be of some use. Many of their deficiencies may be avoided by focusing on relative, not absolute, wealth. While it is possible that the wealthiest citizens were best able to disguise much of their wealth, it is also probable that on any rough categorization - say into four categories of wealth - most of them will still come out at the top. The tax records, then, supplement the use of occupation as a measure of class standing.

Unfortunately, the tax book very rarely listed the address of the property owner. Furthermore, it often listed one owner more than one time, presumably for each piece of property assessed. Two listings of one name could therefore mean two pieces of property owned by one person or it could mean two different owners with the same name, each with one piece of property. I relied on the number of listings in the directory to distinguish between these cases. When guardians had been traced to the directory, the number of listings of their names had been coded. If a name had appeared only once in the directory, all listings of his name in the tax book were assumed to belong to him.

Thus only those whose names were uncommon and for whom an uambiguous directory listing had been found could be traced to the tax book. Many guardians, then, are not included in it. Occupational information available from the census cannot always be related to property value because of these linkage problems. To the extent that property value is important in distinguishing high white collar from low white collar occupations or artisans who are skilled manual workers from those who are petty proprietors, the absence of some tax data will be especially important. Various attempts to work around the problem will have to be attempted.

The information recorded from the tax book was the assessed value of the individual's property.[4] But it is worth noting that of those fathers that could be traced to the tax book, most had no assessed property (67%-70% in different samples). This, of course, is not a problem of missing data, but a statement about property ownership.

B. TRACING THE HIGH SCHOOL SAMPLE TO THE CENSUS MANUSCRIPTS

Most of the high school records did not contain information on ethnicity. The major reason for tracing high school students to the Census manuscripts is to obtain this information. Another, albeit less critical, reason is to obtain information available on the guardian's occupation when school records and directories did not provide it.[5]

The manuscript Census records are not arranged alphabetically. An index exists for the heads of households in the 1900 Census, and all guardians were traced to it. For the 1880 Census an index to heads of households which included a child born after 1870 is available, but this does not, of course, include all the high school students. No index exists for the 1915 Rhode Island State Census. The 1880 sample members not found in the index and the entire 1915 sample, required a special strategy. Every street address falls between two cross streets. These cross streets were found with the help of a special street directory (published in the city directory). Then the appropriate block was located on a map which included enumeration of district boundaries. These boundary lines were obtained from the National Archives. Knowing in what enumeration district an address would fall made it fairly simple to find the address by skimming the street names on each page of the manuscript schedules for that district. If an individual was not found by this method, directories were consulted to determine any recent change of residence, and the alternate addresses were then checked as well. This procedure, of course, is exceedingly painstaking. However, once it becomes routine it is not as gruesome as might be thought at first. Since the high school samples are not very large, it was manageable.

IV. OBTAINING INFORMATION ON SCHOOL ATTAINMENT

A. THE LIMITATIONS OF CENSUS AND SURVEY DATA FOR THE HISTORICAL STUDY OF SCHOOL ATTAINMENT

There are few sources of information which can shed light on the sorts of issues with which this study is concerned. A few communities were the focus of sociological investigation in an earlier period. Similarly government reports, such as the Immigration Commission Reports of 1911, and some School Committee reports will contain relevant material. But in-depth studies will only rarely be available for a particular community or school. And such studies often omit consideration of what we would think of today as key matters - such as the social class origins of students in studies relating school attainment and ethnicity.

One other source that sheds some light on school attainment is the Manuscript schedules from the United States Census and from the censuses conducted by certain states. In the federal censuses three kinds of questions related to education have been asked since 1840. In the 1840-1930 censuses individuals were asked if they could read or write English. While this information is useful for various purposes, it will usually be too coarse an indicator of level of education for most purposes. The second question, included in the Federal Census from 1850 to 1930, asked whether the respondent attended school in the last year. This is a far more valuable bit of information. It is especially useful, of course, when related to the family background information also available in the census. Thus occupation of parents, ethnicity, family structure, residence, community type and so forth may be related to age of school leaving. Unfortunately, age of school leaving is not as clear a measure of grade attainment as one might think at first glance. Two factors explain why this is so. First of all the overwhelming majority of the population left school at roughly the same age - at about 13 through 15 years of age in 1880, for example. And second, children of a surprisingly wide range of ages were included in any given grade. The two factors taken together make it difficult to assess the educational significance of school leaving age.*

Moreover, the census question does not actually state the individual's school leaving age. Rather, it only tells us whether he was still in school in the year of the census. The census information is therefore particularly useful for analyzing the proportion of children in a given social group who were still in school at a certain age. It can, for example, tell us that 30% of the 14 year olds whose parents were born in Ireland were in school in 1880 and 60% of those whose parents were not. But it is much less useful for other purposes, such as relating schooling and later experiences. At what age did an eleven year old reported to have attended school in the Census year actually leave school? And what of a sixteen year old reported as not in school in the census year?

Finally there is, of course, the problem of the veracity of the respondents. After compulsory education laws became known to most families (perhaps in the 1890's in most cities), a family whose child was required to be in school might have responded that he was attending even if the truth was otherwise. Indeed, if the family knew vaguely of the existence of such laws, but not of the specific provisions, they might lie even if their child was in fact already past the age covered in the law.

A third type of question concerning education has been asked by the U.S. Census Bureau since 1940 (and by South Dakota and Iowa Census

---

* See for example the broad range of ages at which children graduated from grammar school in Providence in 1885 (p. 70 below) and the range of grades in which 13 year olds were enrolled in 1917-16 (Table IV-7B, page 64, below.

Bureaus since 1915). It concerns the highest grade of schooling attained by each respondent. How well this question is distinguished from "years spent in school" (regardless of grade attainment) is problematic. But in any case U.S. Census schedules of 1910 and after are closed to researchers. Moreover the Census records do not tell about the family origins of adults. It would therefore be necessary - but exceedingly difficult - to trace an individual 65-70 years old back to the time and place of his childhood to find information about his parents' socioeconomic standing. Indeed, such a trace would be at least as difficult as tracing a child from the Census records of 1880 to school records of the same period (as described below).

It is also important to bear in mind that even the most recent Census data at best tell us one number - highest grade attained. The information one would like - and which can be obtained from school records - is much richer. Curriculum, grades, attendance and possibly "remarks" are included. And even the name of the school is often a tremendously valuable piece of information. The social meaning of a school in a community is a subject for study. It can make a good deal of difference at which school a certain grade of schooling was attained.

In short, in order to study historically the social background of students, and the impact of education on career experiences it is necessary to explore the use of student academic records.

The student academic records themselves, independent of any other sources, can provide useful information. In the register of a high school (a source described in detail below) one can count the number of entrants to each program over a range of years. Moreover, this data can be subdivided by sex in order to study patterns of attendance of men and women by curriculum over time. The register will also provide information on the length of time each student was enrolled and whether he or she graduated.

Such information is useful and easy to obtain. Some material on the social background of the students may also be contained in the school records (for example place of birth, parent's place of birth, perhaps father's occupation, and so forth). In this case the researcher will be fortunate. But I want to emphasize not the use of school records in a vacuum, but the link that can be made between these records and the various sociodemographic records that social historians have mined for the past decade - Census records, tax records, vital records, city directories and others. My point is not that one ought to think about whether such a link is possible, but rather that these social historians have demonstrated that it is indeed possible. Those interested in the social role of schooling ought to explore what aspects of the subject are most worthy of study and whether the school records of a particular community are available which can shed light on them.

I hope, therefore, that the details of my discussion of these records may prove to be of some use not only as background for my own study of Providence but also as an introduction to the kinds of records one is likely to find elsewhere. Although I have examined very briefly the records of a half dozen other Rhode Island towns and those of some New York City schools, I cannot, of course, claim either that the records of other school systems will be identical to those I have seen in Providence, or that the methods I have used will be the best ones in other contexts. But I think there are likely to be a considerable number of similarities in the issues that arise from the use of these records in different places.

A second caveat is also in order. The kind of records I am describing, and even the formulation of the issues I have posed, are most relevant to the latter half of the nineteenth century and after. The nature of educational institutions in earlier periods of American history (for example, the absence of a clear system of grading, undeveloped public secondary schooling and a plethora of academies) suggests both that the questions cannot be mechanically read into earlier periods and that the nature of the older sources will be different. It is probably well worth exploring the use of school records from the early nineteenth century period, but this paper will be less useful for such an exploration than for work in the later period.

*　　　*　　　*

School administrations preserved student academic records long after the students left the schools. They did so not merely to preserve information on school attainment. The records also are often the best available, or the only available, proof that a person is of a given age, or was born or raised in a given area. People came and still come to the schools to certify such facts to the Social Security or Immigration and Naturalization authorites. Some states therefore forbade the destruction of student records for a long time. For example, in Rhode Island their destruction is prohibited for 60 years. Furthermore, having been preserved for so long, they were often preserved much longer. For one thing, if they did not take up a lot of space, or were forgotten in a basement storage area, or were interfiled with other records, it may have required more work to dispose of them than to retain them. And when really old records were rediscovered, school authorities may well have preserved them as quaint, or because they recognized their historical or genealogical value. In Providence, the records of the public high school have survived in apparently incomplete form from its opening in 1843. They are complete from 1877. In New

York City, a school register from the 1830's was discovered in the basement of the Board of Education headquarters. Records of many New York schools, particularly high schools, appear to have survived from the turn of the century and perhaps earlier. Other sources have also been found: midnineteenth century high school records of Beverly, Somerville and Salem, Massachusetts; Massachusetts reform school records of boys and girls from 1840 to 1905; Pittsburgh high school records for the second half of the nineteenth century; records of examinations given to eighth graders in Iowa since the turn of the century;[2] some Boston and Worcester, Massachusetts high school records from the 1890's' the school records of Quincy, Massachusetts from the turn of the century.

Such records, then, do exist for various school systems in the United States. How common their preservation has been will remain unknown until more attempts to locate them have been undertaken. Parenthetically, it might be noted that such a search could be facilitated by an agency such as the National Institute of Education. It would be a relatively inexpensive matter to develop a questionnaire to be sent, (and follow-up calls made) to the school systems in areas that had, for example, a population of 50,000 by 1900, or 25,000 in 1870. The most promising responses could be followed up by site visits to check what was available. While such a project would not turn up the long-forgotten registers in the school department's subbasement, it would spark some searching and would gather together that information which is known to only local school authorities. It would also help prevent further destruction of such records by making school authorities aware of their value.

In any event, the existence of school records is one issue and the permission to use them another. Public school records are no doubt usually, under the control of the city school committee, Catholic school records under the control of the diocese or the parish. Their use, like the use of other records not generally open to the public, raises issues of invasion of privacy and confidentiality. At a time when such concerns are very much in the public (and consequently the public bureaucrat's) eye, it will be necessary to explain quite clearly and fully what one's interests are and what guarantees can be offered that the individual whose records are involved will not be harmed by the research. Similar requirements are made of researchers having anything to do with human subjects by the government and by most universities.

Since such research focuses on aggregates, not on individuals, it should be possible to convince many local administrators that confidentiality can indeed be guaranteed. It is also perfectly feasible to create safeguards to disguise any examples used in a study and to define conditions for the use of the material by others. Whether or not such guarantees will convince a particular official cannot, of course, be predicted. But the experiences of researchers noted above suggest that it often will be possible. A letter of introduction from a well-known academic or state education authority is obviously likely to be an asset. It is also worth keeping in mind the distinction between the right to inspect the available records for a day or two and the right to use the records extensively in a study. If it is possible to obtain the former, it may turn out the records are not worth worrying about, that they are less complete, or extend less far back in time, than expected. In general, negotiations with a public bureaucracy are not necessarily speedy affairs, and one should not expect permission to be given or denied in a day. The request will probably involve school administrators in questions they have not had to face before. Here again the N.I.E. or some comparable authority could be of considerable help with a memo offering suggested guidelines to local school authorities for the use of student records dating from, say, the pre-1940 period or the pre-1900 period.

3. A DESCRIPTION OF THE SCHOOL RECORDS

It will be useful to distinguish between periods in considering the school records that were found in Providence. The records available in 1925 are very different from those available in 1880. And it will also be useful to catalogue the schools according to three types of control: public, Catholic and other private schools. Each type preserved records differently. Finally, it is important to distinguish records by level of schooling. Records of secondary schools were better preserved than records of grammar schools or primary schools. Post-secondary schooling will be considered in a separate section below. Table IV-1 sums up the nature of the records that are available in each period.

1. The Student Academic Records Kept by the Public High Schools.

The student academic records I have used are best examplified by the public high school records. They were recorded first in register books, and later on record cards. The register is a folio-sized, specially printed volume divided into two parts, a register of membership and a register of standing. The first of these includes general background information on each student. The information given was quite consistent over the entire period of the study: name, guardian's name, address, age, previous school, date entered, date left, and a section (usually left blank) for remarks. Occasionally a column was added to provide information on whether a student graduated and if so in what curriculum. The second part of the register, the register of standing, includes each student's attendance record and his grades in each course in each grading period (quarter, semester or year). The register of membership is arranged by entering class: freshmen of 1880 are followed by freshmen of 1881 and then by freshmen of 1882. Within each class the entrants are listed alphabetically. If a student transferred to

Table IV-1 An overview of the most important records available for each level of schooling by period and control of school

Control	Level of Schooling	Period	Available Records
Public	High School	1880-1925	Complete academic records of all who attend
	Grammar School	1880, 1900	Newspaper and alumni lists of grammar school graduates (incomplete in 1880, complete in 1900)
		1915, 1925	Academic records of all who attended (nearly complete in 1915, complete in 1925)
Catholic	High School	1880	Lists of graduates
		1900	Lists of graduates and some academic records
		1915, 1925	Complete academic records of all who attended
	Grammar School	1880	None
	----	1900-1925	Newspaper and alumni lists of grammar school graduates (nearly complete; supplemented by working papers in 1915 and 1925 and also by school census cards in 1925)
Other Private	High School	1880-1925	Lists in catalogues of all who attended the major Providence schools (nearly complete; supplemented by working papers in 1915 and 1925 and also by school census cards in 1925).
	Grammar School	1880-1925	Lists in catalogues of those completing grammar school (some schools only; supplemented by working papers in 1915 and 1925 and also by school census cards in 1925).

NOTE: Very few records of individuals who left school before reaching fifth grade are available, even in the two later periods.

the high school and entered directly into, say, the junior class in 1883, his name was entered at the end of the appropriate list of freshmen (in this example at the end of the list of the freshmen of 1881). The register of standing is arranged by grading period (e.g.: grades for the semester ending June, 1881). And within grading periods it is arranged by class.

To find a student's entire record, then, one would proceed through the grading periods during which the student was enrolled: if he made normal progress through the school he would appear in the freshman class in one grading period, in the sophomore class of the grading period one year later, etc. When a student failed to be promoted he joined the next succeeding class (freshman of 1880 to freshman of 1881) in the register of standing, but was not moved in the register of membership.

The registers were maintained with considerable care. Entries are uniformly neat, almost never crossed out or corrected, and there are virtually no contradictions between the registers of membership and standing (for example, no one appears in the latter but not the former). Some variants on the principles of register organization described here were noted in the records of other cities and even in some Providence registers. But school registers were meant to provide easily accessible information on enrollment. So some ordering principle involving alphabetization and year of entry or class of entry must have been used in most towns. And since enrollment records were probably useful separate from the more cumbersome record of grades, the distinction between a register of membership and a register of standing was probably also a common one.

The register books, and apparently later the cards, coexisted with what was probably a less permanent form of record keeping: the teacher's register book. Each teacher kept such a volume in which to record the attendance and grades of the students. Clearly, it is preferable to use the school-wide register since the same information is recorded in a more convenient form (the student's entire career is in one volume). And since it constituted the permanent record, the school register may have generally been the neater one. But in some schools the school register may have been lost and some or all of the teacher registers retained. A certain amount of patience and care will allow one to extract a considerable amount from the teacher's register as well. In some schools, primary and grammar schools in particular, only teacher registers may have been used.

Readers are probably familiar with student record cards from their own schooling. In Providence the student's name, address and guardian's name were entered at the top and then various kinds of additional information, depending on the period: dates entered and left school (or school system), previous school, cause left school (if not graduation), place of birth, nativity of parents, date of birth, occupation and/or business address of guardian, IQ (only after 1920) and personality ratings (also only after 1920). One card format was prepared by the school system for grammar schools, and separate ones were prepared for each of the four public high schools. The Catholic Diocese used almost identical cards but also listed the parish of the individual. Each semester or year the grade level in which the student is enrolled, his attendance record, and his grades were listed by course name. Grading systems (5 point scale, 10 point scale, A-F, and others) vary with the period, but were either stated explicitly on the card or easy enough to figure out. For all practical purposes the record cards abandon the notion of progress in a "class" in favor of the notion of credits or courses completed.[3] Whether records were kept on cards or in register books also determined how the link between the census sample and the records would be made. Sample members could be traced easily to record cards arranged alphabetically. But tracing them to school registers was another matter.[4]

2. The Catholic High Schools' Student Academic Records

Such, then, is the nature of the public high school records. The other student academic records (listed in Table IV-1) may be treated briefly as variants of the above. No register book was found for the boys' Catholic High School, Lasalle Academy. The school secretary and some of the older teachers remember such a volume, but it has been discarded. By the 1915 and 1925 periods, however, record cards were in use and these remain intact in two record files, one for graduates, one for non-graduates. Luckily, a list of all graduates of the school since its establishment had been prepared from the missing register. Also some teacher registers were available from the 1900 period. The records of the Catholic Girls High School, St. Xavier's Academy, were preserved in a similar form. For the 1915 period sheets indicating the name and grade of each student (arranged by class) were used as well as a list of graduates. For the 1880 period, only a list of graduates was available.[5]

3. Student Academic Records of the Grammar Schools

Finally, student academic records of the public grammar schools are available in the 1915 and 1925 period on cards. A W.P.A. project created a single alphabetical file of all these records, some high school record cards, working papers, and school census cards (described below). The file, which conveniently groups together all surnames that sound alike (and provides cross-references under alternate spellings,) includes about a quarter of a million records. Prior to about 1905, the public grammar school records seem to have consisted of register books which have been lost.[6]

The parochial grammar school records include some teachers registers as well as record cards. But these Catholic records were never interfiled. If a school is still in operation its records are kept there. Others are kept at the Diocesan school office. And in any event for the period of the study the records of most of the parochial schools are incomplete and others non-existent, for even the later sample years. I therefore resolved to limit the inclusion of Catholic elementary school records to lists of grammar school graduates. I supplemented lists from newspapers with lists that were prepared by the schools themselves.[7]

D. Other Records Kept by the Providence School Committee

The sources described above include all the student academic records I was able to use in this study. The reader is reminded of the summary in Table IV-1. But two other kinds of sources provided much additional information on schooling. The first category includes various documents kept by the Providence School Committee other than student academic records. All fortunately were interfiled in the one large file created by the W.P.A. The second category of material constitutes lists of students or of graduates found in various publications.

The school committee kept working papers, school census cards, and high school graduate follow-up cards. Working papers were required under compulsory education and child labor laws. They certified that a youth was legally entitled to work by virtue of his age and/or length of schooling. The record of these certifications was kept by the school department. They include the student's name, date and place of birth, the date of the application, his grade level and school name and his guardian's name and address. They are not necessarily a record of the student's last grade in school, since students seem occasionally to have obtained them before quitting school forever (perhaps in order to obtain summer work). Thus they indicate that a student attained no less than a certain grade by a certain date. Working papers were required by state law from the 1880's but unfortunately are available in Providence for only the 1915 and 1925 periods.

The school census, also an outgrowth of compulsory education legislation, was an attempt to keep track of all youths in the city. It ascertained whether or not they were enrolled in school, and if so in which school and in what grade. It was introduced in the late 1870's, as a once a year enumeration. In 1927 an ongoing census in which records of individuals were periodically updated, was begun. It included all in the age range 5-20. The records of the census effort begun in 1927 have been preserved.[8]

Both the census card and the working papers have the advantage of indicating private as well as public schooling. The school census was meant to cover the entire city. The working papers were limited, it seems, to those who went to work before attaining a certain age. For the 1925 period a follow-up card on public high school graduates is also available. In addition to listing the students' entire academic record, it listed their father's occupation, their birthplace, religious background, occupational aspirations, further schooling and/or early employment.

5. Lists of Students in Newspapers, Catalogues and Alumni Publications

Lists of graduates commonly appeared in the newspapers at the end of the academic year (or at the end of each term when graduation was semi-annual). They include the names of the graduates by school In Providence the graduates of grammar and of high schools are both listed. Often private and Catholic school graduates are also listed. The Diocesan newspaper provides a more complete list of

Catholic school graduates. Lists of grammar school graduates began appearing regularly in the mid-1880's in Providence. Unfortunately, the lists involve serious gaps in the early years. But by a stroke of luck alumni of the two largest public grammar schools each published a booklet listing all graduates of the period.[9]

Finally, the more elite private schools of the city published annual catalogues. These often listed the names of the students, sometimes with their address, and often by grade level. Lists of past graduates were often included also.

In general I was less concerned with the private schools. Their clientele tended to be the small proportion of wealthy children who have been studied in investigations of elite institutions. Furthermore, these schools were not under any central control such as the school committee or the Catholic church, which made searching for their records a much less efficient process: each school required an altogether separate investigation. Finally, these schools were only part of a group of select institutions which included also many schools not located in Providence, to which such children might have been sent. No attempt at completeness of coverage is therefore practical. But the inclusion of the most important of the select schools in Providence will nevertheless reduce the number for whom information on schooling is missing and provide some clues about who attended other select institutions. The inclusion of these schools is especially important in the earliest periods, when a meaningful proportion of those attending secondary school attended the private schools.

6. Use of Census Information on Schooling

Information on educational attainment below the level of secondary schooling will be difficult to come by in 1880. Lists of grammar graduation are incomplete, and in any case only a relatively small number of individuals in fact graduated from grammar school. The only alternative is to rely on the flimsy school attendance data in the census ("attended school in past 12 months?"). As already noted, I have collected a sample of those receiving the least schooling: the 15% of 11 and 12 year olds who were not in school. A second group will constitute those who did not fall into this category but who also failed to graduate from grammar school. A third will consist of grammar school graduates. A fourth will consist of high school entrants (these can be broken down by semesters of high school completed) and a fifth will consist of high school graduates (it may be possible to distinguish in this small group some of those continuing on to college).[10]

7. Post Secondary Schooling

The inclusion of information on post-secondary schooling involves many of the same difficulties as the inclusion of information on private school students. There are innumerable institutions which sample members might have attended, and no easy way to determine which ones they did attend. Luckily there are several factors mitigating the importance of this problem. The first is that the overwhleming majority of sample members could not possibly have reached college. They dropped out of school at a much lower grade level. This is still true, although less so, for the most recent sample, the one from the 1925 period. But - and this is the second mitigating factor - for this sample the school census card and high school graduate follow-up card present information on post-secondary education of high school graduates. Of some 270 high school graduates in the 1925 census sample, information on college attendance is available for about 150. Third, it appears that by 1900 high school yearbooks indicate the graduate's future plans, including college attendance. Since the number of graduates is small, it should be easy to add information from these yearbooks and thus to obtain college attendance data on most of the graduates who attended.[11]

8. Marginal Institutions

This, then, completes the range of educational institutions to which sample members were traced. But there are a variety of more marginal institutions which I made no effort to include: art, music or dancing schools for example. One such marginal institutions is, I believe, more important: the business or commercial school. There were, at any time from 1880 to 1925, several such schools in Providence. Generally, they have either disappeared completely (Scholfield's Commercial College, Max Magnus's School for Shorthand, Spencerian Business College and others), or they continue in one form or another (one as the Katherine Gibbs School, another as Bryant College). I found few clues as to where the records of these schools might be kept. For most, it is hard to believe they have been preserved at all. And the extant catalogues - which are very few - generally do not list enrolled students.

Such institutions offered a very different range of training than the public school "commercial programs" (which were founded partly in response to the success of these business colleges). Some offered courses in traditional disciplines that were useful for business careers - especially grammar, writing, arithmetic. But they also offered a wide range of courses and work programs to provide business skills. These included courses of a year or more but also many others that lasted two to three months.[12] I have no information on those attending such courses. Such information also may well have been missed in the "years of schooling" question in the census (how did people report a three-month bookkeeping or shorthand course taken after leaving grammar school?) Yet such schools may have been enormously important for a critical group of individuals: the children of blue collar, often immigrant parents on their way into clerical occupations, and sometimes into business management. We just do not know anything about their clientele. And the inability to probe this question seems to me a more unfortunate omission - given the concern with the mechanisms of social mobility for those starting near the bottom - than the omission of a number of private boarding schools.

C. LINKING SAMPLE MEMBERS TO SCHOOL RECORDS

The task of matching sample members to the records described above required first of all a way of comparing the two. Two categories of records should be distinguished: public school records alphabetically arranged in card files at the school department and all other records. It was relatively straightforward, if time consuming, to trace sample members to the card files. And since no records of the 1880 period were in these card files, the 1880 sample members, nearly 3000 in number, could be excluded from that trace. On the other hand, to trace 12,000 sample members to each of 100 or so registers and lists that comprised the other category of records would clearly have been impossible. Instead the names of students in these sources were copied, along with any information that would identify them: address, guardian's name, year of graduation or entrance. The resultant list of student names was then alphabetized and all sample members could be traced to it.

In copying the names of students from the school sources, I was, of course, not interested in every student who attended between 1880 and 1925. Rather, I was only interested in those who by virtue of their age could conceivably have fallen into my samples. For example, the 1900 census sample includes boys who were 13-16 in 1900. A boy listed in the school records as 18 in 1902 might have fallen into the sample. But a boy listed as 12 in 1902 could not have, unless there was a significant error in recording his age. Similarly, although the published list of grammar school graduates do not provide ages, it is possible to say with a high degree of probability in what range of years boys 13-16 in 1900 would have graduated from grammar school. The school committee's Annual Reports often listed the age distribution of graduates, or an age-grade report. Using this information for the 1900 period, I included all the names of graduates from 1897 through 1903. Many of these graduates would be older or younger than the sample members. But very few of the sample members would graduate in other years. In the case of lists of private high school entrants I used the same approach.

Once alphabetical lists of students had been created, sample members were traced to these lists in one of two ways. At first I tried to exploit the sophisticated computer programs which can link records from two machine-readable files. I therefore made machine readable the information I had collected both on the sample members and on the students culled from the school records. The computer program (I used a version kindly provided by Professor Theodore Hershberg of the Philadelphia Social History Project) could do two things.

First, it would assign each name a short alphanumeric "soundex" code based on its sound. Consonants, or combinations of consonants, that sound alike (or look alike) would be assigned the same number. Similar names will thus be clustered together despite minor variations in spelling. The second purpose of the program is to compare all individuals assigned the same soundex code in one file, to all who received that code in the other file (e.g.: to compare all sample members coded A135J200 to all grammar school graduates who received that code). The comparison must of course be made on variables that are the same in each file: age, place of birth, etc. But I had too few such variables available. And two important ones - the sample member's address and guardian's name - had not been put into machine-readable form for the sample members. In short, I could not exploit that part of the program which would have actually linked sample members and students. Instead I used a simpler program. All the students assigned a certain soundex code were printed alongside the names of the sample members assigned the same code. Then the comparisons on each variable were made by hand.[13] The computer, in other words, saved the time and effort needed to alphabetize the students' names, and to trace each sample member to the alphabetized list. The new school information derived from the linkage work was then added to the machine readable information already available on the sample member. But the process proved too complicated. The advantages of having the computer do the tracing had been outweighed by the complexity of shifting from machine-read computer tapes to comparisons made by hand and back to the computer to update the tapes.[14]

Therefore, when it came time a year later to process the second half of the sample members, I changed the procedure. My revised procedure was simply to let the computer alphabetize the names of students. I postponed making information on the sample members machine readable. The sample members were then traced to an alphabetized printout of students much as they would be traced to a city directory. The information on schooling was then copied onto the sample members' forms and when the forms were later keypunched, the new information on schooling was simply keyed along with the rest.

Everything I have described concerning linking the sample members and the school records up to this point has had to do with the purely logistic problem of how to make the comparison most efficiently. The second linkage problem was the more familiar one: by what criteria to determine whether a sample member and a student listed in the school records with the same name were indeed the same individual. When the linkage involved school records (rather than merely names from newspaper or catalogue lists), three bits of information were provided which usually made the matching process very simple. These are the age, address and guardian of the student. If the student's age was within a range of three years of the sample member's in the same year, and if either the address or the guardian's name matched, I assumed the student and sample member were one.

Actually, a very high proportion, probably well over 90% of all those "successfully matched," were no more than a year apart in age. When the name and age matched but neither the guardian's name nor the address matched, or when a sample member could be matched by these rules to more than one record or in the event of any other difficulty, the research assistants who used these rules showed the case to me. With the aid of city directories it was virtually always possible to resolve satisfactorily whether the record referred to a particular sample member or not.

When the sources consisted of only lists of names from newspapers or catalogues the task of linkage was much more difficult. Several criteria were, however, available. One was age. Sample members were traced to lists of students covering a range of years. But not all sample members could have been enrolle in every one of these years. For example, the 1900 period sample members were traced to lists of graduates of 1897-1903. But the younger sample members could only have graduated in the later years, and the older ones in the earlier years of this range. A second criterion was geographic area. Schools were restricted to residents of school districts and these districts followed ward lines. Since wards were indicated in the census records it was possible to determine whether the sample member and the graduate of a grammar school who had the same name came from the same part of town. If they did not, it was clearly less probable that they were the same individual.[15] The third criterion was the commonness of the name of the sample member. This had been coded when the men were traced forward in time to the city directories, and the women to marriage records. If a name was uncommon, it was far more likely that sample member and graduate were one. Similarly, the same middle initial in each name increased this probability.[16]

All of these criteria involve probabilities, not certainties. They strengthen or weaken a match based on name only, but they do not definitely exclude or confirm it. Each criterion has a precise weight relative to the others which, taken together, would produce the best guess of whether the match is correct. The uncommonness of the name may be two and one-third times as important as a contradictory middle initial, for example. To derive such weights empirically would have been a large project in itself. Luckily, I had few cases that actually had to be linked in this way, and I therefore devised instead ad hoc weights based on my work with the samples. I doubt that this procedure made much difference for the results. But because the nature of the matching to mere lists of names is far less certain than to full academic records, I limited the use of such lists as much as possible. Specifically, when high school entrance was already known from school records, I did not try to obtain information on grammar school graduation from lists. Moreover, I divided the sample members for whom linkage scores had been calculated into five categories of increasing probability. Schooling information was recorded only when the score fell in one of the two most probable categories. These constituted about 40% of all pairs of sample members and students with the same name. For the three weakest categories only the level of schooling involved (grammar, secondary, post secondary) and the strength of the linkage was recorded. It will thus be possible to exclude these individuals from analysis or to classify them first in one category of educational attainment and then in another to see if the results are affected.

The Information Recorded on Schooling

It will be clear from what has already been said, that the nature of information available on each sample member's schooling was determined by the kind of source available. School records differ from working papers, which in turn differ from newspaper lists. In addition the school records themselves were somewhat different in each period and in each institution.

Nevertheless, it proved feasible to develop a form for recording information on schooling which, I think, involved very little distortion or loss of the material that was available (see attached form). The material was divided into four parts: the student's social background, post-secondary education, secondary education, and lower levels of education.[17]

D. THE COMPLETENESS OF THE RECORDS AND THE ACCURACY OF THE LINKAGE PROCEDURES

I have assessed the completeness of the data collected on educational attainment as carefully and painstakingly as possible. These data are not only at the heart of the study, they are also by far the most complicated to collect and the ones for which population parameters (against which to check the adequacy of the results) are least readily available. Moreover, records for education below certain levels of schooling were not available. It is therefore important to determine whether the absence of school records implies that an individual received relatively little schooling, or that his records for a higher level of educational attainment were simply missing, or finally that my use of them was somehow flawed.

Table IV-1 sums up the nature of the sources which were covered and the levels of schooling included. Basically, I sought completeness of information about the public and Catholic grammar school graduates, and high school students, and about the students at most of the important select secondary schools in Providence. Also, in 1915 and 1925 the records of all public grammar school students, graduates and non-graduates, were available. Some other records were found but the educational levels mentioned are the ones for which I hoped the data could be complete. Omissions could come from gaps in the sources themselves, or from random or systematic errors in linkage procedures. And of course even if none of these problems exist, the sample might be unrepresentative of the population. Finally, errors would occur because of the omission of records from institutions I could not even try to include: primary schools and some select private schools in and especially outside Providence.[18]

In assessing the completeness of the data listed in Table IV-1 I made a number of preliminary checks before collecting the school records. At the most basic level I had to be sure that I was in fact aware of all the schools that had existed.

The Annual Reports, Manuals and other publications of the Providence school committee provide this information on the public schools. Histories of the Catholic schools (particularly a thorough doctoral dissertation of 1957) and the annually published Catholic Directory were important for the Catholic schools. Also the author of the most detailed history of Rhode Island, Charles Carroll, came to his subject from studies of the educational history of Rhode Island. Several long chapters of his history, as well as his other work, provide a guide to the evolution of the schools. Other histories were, of course, also consulted. And catalogues of all the major private schools were available. Finally, an invaluable record of the schools between 1893 and 1920 are the tables prepared annually by the Rhode Island Commissioner of Education in his Annual Report. They include for every non-public school in the state, and for every public high school, enrollment by sex, number of graduates, number of students over age 15, and various other information. When a school failed to report (a violation of state law) the school's name was still listed. Thus these tables provide an important guide to non-public education in Providence.[19]

I also checked the records themselves in various ways to be sure they constituted complete collections. For example, I verified that they included both boys and girls, graduates and non-graduates, all parts of the alphabet, and (in the case of high schools) all known curricula, and all relevant years of entry. Finally for every public grammar school and high school in the 1915 and 1925 periods a number of graduates were selected from the newspaper lists and traced to the record cards. This test revealed that the records of 5 grammar schools were imperfect in the 1915 period. They were therefore supplemented by the use of newspaper listings of graduates. Otherwise, these preliminary tests revealed no gaps in those academic records thought to be complete. (See Table IV-1, page 32 above).

But of course if a relatively small percentage of the records were missing, these tests would not reveal the gaps. And in any case the records of sample members might have been missed in the tracing process. The more important tests, then, involve the analysis of the data actually gathered on sample members. While there are any number of such checks that could theoretically be made, most would prove inconclusive because of the small number of sample members involved (such as the number graduating from each grammar school). I have instead approached the question in several ways.

1) I have tried to estimate the number of individuals in my census samples who would be expected to have graduated from public or Catholic grammar schools on the basis of available published information. Then I compared these expected numbers to the actual numbers of such individuals in the samples. 2) I then made a similar comparison of the expected and actual number of sample members attending select private schools. 3) I compared the number of sample members estimated to have attended certain high schools to the number who actually did. This test differs from the first two not only in the level of schooling involved but also in the confidence one can place in the estimate. The number of high school students in the sampled age range in these high schools is known from the register books, so the estimate should be quite accurate. 4) Next I

examined the individuals on whom no school information is available. I explored, with the help of the calculations made in the preceeding tests, how many of these individuals are likely to have reached or completed grammar school, not withstanding the fact that I found no evidence of that attainment. 5) And finally, again with the help of calculations made earlier, I estimated the number of sample members who never reached the fifth grade and compared that estimate to estimates from published data as a kind of independent check on the accuracy of the entire set of procedures described here.

1) Public and Catholic Grammar School Students in the Samples: Expected vs. Actual Numbers. This approach rests heavily on estimates of the number of grammar school graduates expected to fall into the sample. The details of the estimating procedure will be found in the note at the end of this section. But in general it involves 1) ascertaining the number of graduates in each year in which sample members might have graduated, 2) determining the proportion of these graduates who were the same age as the sample members selected from the census records, 3) determining the proportion of graduates of each sex, 4) adjusting the figures slightly to take account of population turnover (some graduates could not have fallen into the sample because they were not living in Providence in the census year and some sample members moved out of town and graduated elsewhere), and 5) multiplying the result by the sampling ratio (e.g.: by 1/6 in 1900). The result is the number of graduates expected to fall into the sample. The difference between it and the number who actually were found may be due to sampling as well as to incomplete records or faulty tracing. And it is important to bear in mind that the procedure rests on the accuracy of the initial assumptions about numbers of graduates, proportions in the sampled age range, the sex ratio of the graduates and the extent of population turnover. Virtually every one of these

magnitudes rests on estimation and interpolation--that is to say, on some guess work from incomplete published and unpublished data. It is almost certain that they involve some error. On the other hand, while I would not put much confidence in the exact magnitude of any result in the following tables, most are probably reasonable approximations. If, for example, the expected number of graduates is estimated at 5 percent of the sample, it may be as low as 3 percent or as high as 7 percent. Perhaps it is even 10 percent. But it is very probably not as high as 20 percent.

And, as the figures in Table IV-2 indicate, in no case is the difference between the expected number and the actual number more than 6% of the entire sample. The total proportion of sample members who have been misclassified on educational attainment is therefore low. But two points should be considered. The first is that all the errors due to missing records or to faulty tracing are of one type: they attribute to sample members less education than really received. Thus the errors do not, unfortunately, cancel each other out; they lead to a downward bias. That bias, luckily, is slight because of the size of the error in comparison to the size of the sample. The second point, however, is that the missing records affect certain subgroups of sample members more than others. In particular, of course, Catholic school records affect Catholics most, and particularly Irish and French Canadian Catholics (the Catholic groups who enrolled the largest proportions of their young in these schools). On the other hand the small number of records missed, and the large number of Catholics suggest that even for this group only a relatively small proportion of errors will result.[20]

Table IV-2. A Comparison of the Expected and Actual Number of Graduates of Public and Catholic Grammar Schools in Each Sample and of Entrants to these Schools in 1915 and 1925 Samples

Period	Sample (N)	Type of School	Level of Grammar School Attainment	Number of Such Students in Sample Expected[1]	Actual	Difference	Difference Expressed as a Percentage of the Entire Sample
1880	Male (1223)	public	graduates	160	97	63	5%
		Catholic	graduates	12-35	1	11-34	1%-3%
	Female (898)	public	graduates	160	111	49	5%
		Catholic	graduates	13-38	1	12-37	1%-4%
1900	Male (914)	public	graduates	191	174	17	2%
		Catholic	graduates	53	16	37	4%
1915	Male (915)	public	graduates	363	335	28	3%
			other entrants	312	278	34	4%
		Catholic	graduates	56	55	1	0%
			other entrants	58	22	36	4%
	Female (903)	public	graduates	426	376	50	6%
			other entrants	236	216	20	2%
		Catholic	graduates	85	71	14	2%
			other entrants	63	57	6	1%
1925	Male (1090)	public	graduates	519	507	12	1%
			other entrants	255	242	13	1%
		Catholic	graduates	102	106	(+4)	0%
			other entrants	75	23	52	5%

1. See Note at end of Section IV of the text.

But a further point should also be noted. The question posed in the preceding paragraphs, and in Table IV-2, is how large a proportion of the entire sample (or of an ethnic or religious group within the sample) has been misclassified on the schooling variables. Another way to approach the problem is to ask what proportion of individuals characterized by a given level of education (e.g. grammar school graduation) have been misclassified. For example, the first line of the table indicates that 160 public grammar school graduates might have been expected to fall into the sample, and 97 actually did. The difference (160-97=63) amounts to 5% of the entire sample, a fairly low figure (and most in the table are lower). But if the question is the proportion misclassified of all individuals who graduated from public grammar schools, the estimate is $63/_{160}$, or a very substantial 39%. This conclusion, if reliable, should certainly mean that one cannot generalize with as much confidence about the characteristics of the subpopulation who completed public grammar schools because an estimated 39% of that group has somehow been missed.

But two points should be noted in this connection. The first is that the sample of public high school entrants of 1880, 1900 and 1915 can be exploited for much better information, if not on grammar school graduates at least on high school entrants. Second, I simply do not believe that the calculations by which I estimated the expected number of graduates are accurate enough to be used to judge the proportion of graduates not found. It is one thing to say that 160 were expected, 97 found, consequently an estimated 63 missed, and that, 63 ÷ by 1223 is roughly 5%. Divided by 1223, it matters little if the estimated 63 is in fact 43 or 83. The proportion misclassified (5%) would only vary by ± 1.6%. But if the denominatior in the calculation is not the sample size (1223) but the expected number of graduates (160), any error in the estimations is very much greater because the denominator is very much smaller (160 vs. 1223). The reader should recall my warning that five sets of estimates were piled up to obtain the expected number of graduates. And the reader who wants a better feel for these estimates is again referred to the appendix to this chapter.

In sum, the crudeness of the estimates will be refined if additional relevant information is found. But given the data at hand these crude estimates can be used with some confidence to estimate the proportion of the sample misclassified but not the proportion of grammar school graduates misclassified. This point should also be borne in mind in reading the rest of the estimates in this chapter.

2) Private School Students in the Samples: Expected vs. Actual Numbers. The method of estimating enrollment in the select private schools was different since information on the number of their graduates would have been hard to piece together. Instead, I relied on the school census reports on the number of individuals of school age enrolled in each type of school (public, Catholic or private). The proportion of all children in private schools was compared to the proportion of sample members whose last school was a private school. Both figures count each individual only once and examine only one point in the individual's school experience (the census is limited to school attendance at a given date, the sample data to a given point in school career). But of course they are not strictly comparable measures. Nevertheless, the proportion of the population attending these schools was so low that the estimates of error cannot be far off (Table IV-3).

These estimates probably include students at boarding schools outside the city. There is no mention of the exclusion of such students from the school census, and since the purpose of that census was related to compulsory education laws, it would have been relevant to include these children. Nor were these children likely to come from families that would have reason to misinform the census taker, or be missed by him.

3) Students in Certain High Schools in the Sample: Expected vs. Actual Numbers. The third comparison (Table IV-4) involves those high schools for which an accurate count of entrants in the sampled age range is available. This count comes either from the lists of students copied from school registers for linkage purposes (see above page 42) or from the sample of high school students selected from the same source. In order to obtain the number expected to fall into the census sample it was necessary only to omit

Table IV-3. A Comparison of the Expected Number of Sample Members Enrolled in (non-Catholic) Private Schools and the Number of Sample Members whose Last Recorded Schooling was in these Schools.

Period	Sample	Number expected to enroll in private schools	Number whose last recorded schooling was in these schools[1]	Difference	Difference expressed as a percentage of the entire sample
1880	Male	62	31	31	3%
	Female	46	9	37	4%
1900	Male	16	3	13	1%
1915	Male	10	8	2	0%
	Female	10	5	5	1%
1925	Male	16	8	8	1%

---

1. The number expected to enroll is based on the proportion of all school age children who attended private schools. The proportions are those found in the School Censuses of 1880 (5.1%), 1900 (1.8%), 1915 (1.1%), and 1925 (1.5%). They are reported in the Providence School Committee, Reports for those years.

Table IV-4. A Comparison of the Expected and Actual Numbers of Sample Members who Entered Selected High Schools

Period	Sample	Number who entered selected high schools[1] Expected	Actual	Difference	Difference expressed as a percentage of the entire sample
1880	Male	132	95	37	3%
	Female	131	107	24	3%
1900	Male	116	90	26	3%
1915	Male	119	128	9	1%
	Female	105	109	4	0%
1925	Male	124	118	6	1%

---

1. These high schools were selected because the names of all entrants in the sampled age range had already been copied. They had been copied for the purposes of record linkage or for inclusion in the high school samples. A small adjustment in the expected number was made to take into account the effects of population turnover (see the note at the end of Section IV of the text, part D).

students who lived out of town(identified by school record) address and adjust the figures slightly for the effects of population turnover.

The figures in Table IV-4 are, then, a kind of check on the first comparison (in Table IV-2). They involve a more narrowly defined group but a more accurate estimate. Of course, since the lists were copied from school records they do not test the completeness of those same records. But they do test the effects of errors in tracing. And once again the difference between the expected and the actual amounts are small, a trivial proportion of the entire population.[21]

4) *Sample Members in Grammar School Whose Records were Missed: Proportion they Constitute among Sample Members Lacking School Records*. Central to the foregoing argument is the notion that the available school records preserve information on individuals who reached a certain level of schooling, but not about those who did not. In the first two periods an individual had to *complete* grammar school to turn up in the records. In the last two periods he only had to *enter* grammar school in order to turn up there (See Table IV-1). Therefore, if the records, and the use made of them, had been perfect, one could assume that sample members for whom *no* school record was found had not reached these levels of education. But probably neither the preservation nor the use of the records was in fact perfect. As a result, some individuals in the two earlier samples graduated from grammar school, and some individuals in the two later samples at least entered grammar school even if no record of these attainments was found. These individuals would therefore be misclassified if all sample members without extant school records are assigned to the lowest category of educational attainment.

The preceeding tables presented evidence to show that the misclassified individuals probably constituted only very small proportions of the entire sample. These misclassified individuals will now be considered in terms of the proportion they constitute among the whole group of sample members for whom no school records were found. If they constitute only a small proportion of the group without school records we can safely assign that whole group to the lowest category of educational attainment.[22] In fact, the error due to the misclassified individuals constitutes only about a tenth of all without school records in the two earlier periods.[*] Hence for many purposes it will indeed be possible to ignore any error caused by the misclassified individuals. In the two later periods, they constitute a higher proportion. But even then most of the error that does occur, at least in the samples of men, ensues from classifying those who received some grammar schooling (but did not graduate) as having received no grammar schooling. This is not a great distortion (as for example classifying high school graduates with those receiving less than a fifth grade education would be). As in the earlier periods, only a relatively small proportion of those classified into the lowest category of educational attainment in fact graduated from grammar school (12% -18%). However, in the 1915 sample of girls the comparable proportion is 53%. This figure is surprisingly high given the proportions in the other five samples, and given the fact that the records and methods used were the same as for the other 1915 and 1925 samples. Perhaps the error in the estimates is for some reason larger. But in any event, if subsequent study does not revise this estimate, greater care will be required in

---

[*] I.e.: 100% less the proportions in the last two columns of Table IV-5.

Table IV-5. The Estimated Educational Attainment of Sample Members for Whom no Record of Schooling was Found

Period	Sample	Number of sample members for whom no record of schooling was found	The estimated number included at left, and their proportion of all included there, who:										The remaining number of sample members for whom no record of schooling was found and their proportion of all not found (first two periods: did not complete grammar course; last two periods: did not enter it)		The proportion of sample members for whom no record of schooling was found who are assumed not to have completed grammar school
			completed a public grammar school(1)		completed a Catholic grammar school(1)		entered but failed to complete a public grammar school(1)		entered but failed to complete a Catholic grammar school(1)		attended a non-Catholic private grammar school(2)				
			N	(%)	N	(%)	N	(%)	N	(%)	N	(%)	N	(%)	%
							(1915 and 1925 only)								
1880	Male	1086	63	(6)	11-34	(1-3)					31	(3)	981-958	(90-88)	90-88
	Female	765	49	(6)	12-37	(2-5)					37	(5)	667-642	(87-84)	87-84
1900	Male	709	17	(2)	37	(5)					13	(2)	642	(91)	91
1915	Male	171	28	(16)	1	(1)	34	(20)	36	(21)	2	(1)	71	(42)	82
	Female	134	50	(38)	14	(11)	20	(15)	6	(4)	5	(4)	39	(29)	47
1925	Male	160	12	(8)	-	(0)	13	(8)	52	(33)	8	(5)	75	(47)	88

Notes to table: 1. from Table IV-2
2. from Table IV-3

treating the group without school records in this sample.

In connection with all the samples, it should also be recalled that the high school records are much more complete than the grammar school records, so the individuals assumed by these estimates to have been misclassified are not likely to have continued on to high school (since no high school record was found for them).

Once again it is important to remember that the estimates are very rough. Moreover, any errors are likely to distort the conclusions just discussed (those derived from Table IV-5) more than earlier ones. If an estimate, for example an estimate of the number of grammar school graduates, is off by 40, the error amounts to a mere 4% of an entire sample of 1,000. But the same error constitutes a much larger proportion of the group for whom no school records were found. In 1925, for example, there are only 160 such individuals. An error of 40 would amount to 25% of this group.

5) Sample Members Who Never Reached Grammar School: Estimates Based on the Preceeding Calculations and Comparable Ones from Published Data. Finally, consider again the group for whom no school records were found, who are assumed to have received little schooling, and the misclassified individuals among them whose records were somehow missed. I just discussed the low proportion which the latter constitute among the former in most of the samples. Now, if the individuals assumed to be misclassified are subtracted from the total number without school records, the remainder are sample members assumed to have indeed received the lowest levels of education. This subtraction is carried out in Table IV-5, and the numbers remaining appear in the next to last column. In the 1915 and 1925 periods these sample members are thought to have received only a primary school education (grades 1-4). For a few other sample members some school records were available which indicate that

Table IV-6. The Estimated Number of Sample Members Who Did Not Reach the Fifth Grade: 1915 and 1925 Samples

Period	Sample	The number for whom no school records were found, less those assumed to have reached the fifth grade (from Table IV-5)	The number whose last recorded schooling was:		The total number of sample members assumed not to have reached the fifth grade	
			in grades 1-4	in an ungraded classroom	N.	% of entire sample
1915	males	71	26	11	108	11.8
	females	39	18	10	67	7.4
1925	males	75	8	25	108	9.9

they too never reached grammar school. These two groups of sample members considered together constitute 11.8% of the boys and 7.4% of the girls in 1915 and 9.9% of the boys in 1925 (Table IV-6). These figures, then, are an estimate of the proportion of all sample members in the two later periods who never reached fifth grade.

This estimate is in turn based on the earlier ones (used in the subtraction in Table IV-5). Hence to gauge its accuracy is to gauge theirs to some extent as well. And it so happens that it is in fact possible to compare these estimates of the proportion never reaching fifth grade in the samples to some published data. The Providence School Committee published an age-grade report in its annual Report beginning in 1917. The age-grade report indicates the number of pupils of each age and the grade in which they were enrolled. It is possible, on the basis of this report, to estimate the number who dropped out before reaching grammar school.

Comparing enrollments in successive grades (e.g.: number in grade 2 divided by number in grade 1) appears to lead to very serious errors, as explained in the notes.[23] But a good deal can be learned by focusing instead on the enrollment at the highest age for which schooling was compulsory in Rhode Island in these years, namely 13. Of course, if many students left school before that age despite the laws, the enrollment figures would not be representative of all 13 year olds. But in fact, very few students left school before age 14. This is evident from Table IV-7A. It presents the enrollment by age for the years 1917-1919 (the first years for which the age-grade report is available) and for 1924-1926 (the years most relevant to the 1925 sample). The enrollment at age 13 is a very high proportion of the enrollment at age 7, and presumably all 7 year olds were in school. Moreover, the small differences between the number of 7 year olds and the number of 13 year olds are not due exclusively, perhaps not even primarily, to dropouts but rather to insignificant differences in the number of children at different ages. Thus, the decline between 7 and 13 is not an orderly one, advancing with each year of age. Nor is there any clear pattern of decline that holds for all or even most of the years 1917-19, 1924-26. There were, for example, considerably fewer 11 year olds than 10 year olds in 1917. But this does not indicate that many children tended to drop out at age 11, for there is no such pattern in subsequent years. Thus, while a few children may have left the public schools before the age of 14, it seems reasonable to assume that the overwhelming majority--the figures suggest well over 90 percent--did not.

Therefore, the grade attainment of the 13 year olds may be taken as a minimum estimate of the highest grade attained by each pupil. Many would, of course, remain in school until they were 15 or 16 or even older. Others would leave immediately upon turning 14. But the highest grade attained by each was no lower than the grade in which they were enrolled at age 13.

Table IV-7B shows that in each of the years except 1917, the number who had not reached grammar school constituted only 12%-14% of the 13 year olds. Nearly all of these children were in ungraded classes of various kinds (such as classes for disciplinary cases) rather than in the regular primary grades. In 1917 the number who had not reached the grammar schools is slightly lower (8.5% of the age group). The difference may be due to the expansion of the ungraded programs beginning in 1918.*

---

* The very low proportions who had not reached 5th grade suggest that even if a small number had dropped out, and even if they came disproportionately from the primary grades, they would not affect the results in Table IV-7B much. If, for example, the figures for 1925 ignore 5 percent of the age group, and if fully half of these 5 percent had left school while still in the primary grades, the proportion of 13 year olds who had not yet reached grammar school would only rise from the 12.3 percent stated in Table IV-7B to 14.2 percent. Moreover, while some 13 year olds had in fact already dropped out, it is also true that some still in the primary or ungraded classes might have eventually reached grammar school. For this reason also, the proportions presented in Table IV-7B probably are not underestimates.

Table IV-7A. Enrollment by Age in Several Years

Enrollment of each age group expressed as a percentage of enrollment of 7 year olds in:

Sept. AGE	1917 %	1918 %	1919 %	1924 %	1925* %	1926* %
7 (N=)	100 (3488)	100 (3412)	100 (3468)	100 (3814)	100 (3798)	100 (3620)
8	100	102	99	95	95	112
9	98	101	101	94	94	109
10	96	101	99	97	90	107
11	89	99	100	98	100	108
12	88	93	99	94	98	113
13 (N=)	86 (3010)	90 (3061)	92 (3205)	94 (3602)	95 (3613)	115 (3748)
14	51	56	59	79	92	106
15	33	32	35	47	58	65

* The compulsory education laws required attendance through age 14 in these years.

Table IV-7B. The Grade Level of 13 Year Olds in Several Years

Proportion at each grade level in:

Sept. Grade Level	1917 %	1918 %	1919 %	1924 %	1925 %	1926 %
Un-graded	6.0	12.0	10.4	9.9	10.8	11.7
Primary	2.5	1.8	2.6	1.8	1.5	1.1
Grammar	77.3	76.2	75.2	74.0	72.1	73.1
High school	13.7	10.0	11.9	14.2	15.6	14.6
TOTAL	100	100	100	100	100	100
(N=)	(3010)	(3061)	(3205)	(3605)	(3613)	(3748)

Source: The age-grade reports published in the Providence School Committee, *Reports* for the years indicated.

How well, then, do the proportions not reaching grammar school in the age-grade reports correspond to the proportions estimated from the sample data of 1915 and 1925 (summed up in Table IV-6)? The answer is, they correspond very well. The sample data and estimating procedures led to estimates of 7.4%, 9.9% and 11.8% who had not reached grammar school in the different samples. In the age-grade reports the comparable proportions for 13 year olds were 8.5% in one year and 12%-14% in the others. Combining the estimates for male and female sample members not reaching grammar school in the 1915 period provides a figure which is more comparable to the age-grade data, since the latter is not broken down by sex. The figure for the two sexes taken together is 9.6% in the sample data.[24]

It is possible that any of these estimates from the samples are a few points off--those for the 1915 girls (7.4%) and the 1925 boys (9.9%) might well be a bit low, the one for the 1915 boys (11.8%) a bit high. In the absence of age-grade reports for the years preceeding 1917 and in the absence of the age-grade data by sex, it is hard to be sure.[25] But in any case, only a few percentage points (less than 5) separate the estimates. The major point is that the magnitudes--7.4% - 11.8% on the one hand and 8.5% - 13.8% on the other--are impressively close.

* * *

To sum up, this exercise in estimation began by attempting to state how many sample members might have been expected to complete public or Catholic grammar schools or to have attended a private school. The expected numbers were compared to the number actually found to have done so in each sample. The differences were always only a small proportion of the sample. The same result was obtained when the number of sample members expected to have attended certain high schools was compared to the number who actually did. Then attention focused on the small differences between the actual and expected numbers of grammar school graduates and private school students in the samples. These differences were assumed to have been caused by the presence of sample members who had received a given level of education but whose records had somehow been missed. They constituted relatively small proportions of all sample members without school records. Generally, then, when no school record was found the educational attainment of the sample member is indeed low. Finally, the number whose school records were thought to have been missed were subtracted from the entire number of sample members without school records. The number remaining were assumed to have received less education than the records available for that period would capture. For the last two periods this meant the sample members would not have reached the grammar school grades. School records had also been found for a small number of other sample members indicating that they too had not reached grammar school. These two groups were combined. The proportion they constituted of the entire sample was compared to the proportions of 13 year olds who had not reached grammar school in the published age-grade reports for 1917-19, 1924-26. The results based on the sample data and the ones derived from these reports were quite close. Thus, one method of estimating the completeness of the educational data - using the difference between actual and expected numbers - was confirmed by an altogether independent method - using the proportions below fifth grade in the age-grade reports.

The quality of the sample data on schooling will have to be considered and reconsidered as the research progresses. And it cannot be too strongly emphasized that all of these estimates leave plenty of room for error. But the consistency of the results when using different methods and when applying these methods to several independent samples, is encouraging. On the whole, then, it appears that the number, and especially the proportion of sample members for whom the absence of pertinent educational data has led to their misclassification is very low and should not seriously affect the analysis.

NOTE TO SECTION IV: CALCULATING THE EXPECTED NUMBER OF GRAMMAR SCHOOL STUDENTS IN TABLE IV-2.

The calculation is based on

A) the number of graduates in each year in which sample members would have graduated (for example 1897-1904 for the 1900 sample);

B) the proportion of all graduates in each of these years who were in the sampled age cohort (e.g.: the proportion that would have been 13-16 in 1900, 14-17 in 1901, etc.).

C) the proportion of graduates of each sex;

D) the proportion of graduates who were living in the city at the time of the Census and the proportion of sample members who, although they lived in Providence at the Census date, actually graduated from grammar school elsewhere.

E) the sampling ratio (e.g. 1/6 in 1900).

The expected number of graduates is the product of these factors (AxBxCxDxE). Each of the first four had to be estimated in most cases. The sources for, and logic of, the estimates follow.

A. The Total Number of Graduates

--PUBLIC SCHOOLS

1915, 1925: number of students promoted in highest grammar school grade (Providence School Committee, *Reports*, 1912-1929).

1900: number entering highest grade of grammar school X .90 (to reflect attrition in highest grade). Numbers entering available only for 1898-1900 and 1902 (*Reports*, 1903-43). Average of these years used for other years. Attrition in the highest grade equaled 11% in 1905-7 (*Reports*, 1407, 40).

1880: a) complete lists of graduates available 1881, 1884, 1885 in newspapers.
b) lists of the two largest schools available for all years.
c) numbers entering high school in each year available in school register.

Number of graduates estimated in two ways:

$$= \frac{\text{a for 1881, 1884, 1885}}{\text{b for 1881, 1884, 1885}} \text{ X b for other years}$$

and

$$= \frac{\text{a for 1881, 1884, 1885}}{\text{c for 1881, 1884, 1885}} \text{ X c for other years}$$

The two results were nearly identical.

--CATHOLIC SCHOOLS

1900, 1915: number of graduates listed annually in R.I. Commissioner of Education, Annual Reports, 1897-1919.

1925: numbers for 1926, 1927, 1929 found in diocesan records; other years were estimated from these and from Ibid, 1920,21.

1880: (ratio of $\frac{\text{catholic}}{\text{public}}$ enrollment in school census each year)

X (estimated number of public school graduates each year)

= maximum figure. Minimum figure = 1/3 Maximum (a guess).

The school census figures were reported annually in the School Committee Report. The ratio is very close to that obtained by comparing public enrollment (from Reports) and Catholic enrollment (from Catholic Directory) in 1881, the one year in which the latter information is complete.

The ratio almost surely overestimates the number of Catholic school graduates. In 1880 the Catholic schools were of two types: parochial schools and academies. The bulk of the enrollment was in the former. But the parochial schools no doubt enrolled mainly primary grade students. The academies all seem to have enrolled grade school level students. Indeed, some appear to have differed from Catholic grade schools only in name and pretentions. Most also were not supported by a single parish and they charged tuition. The relation between the two types of institutions is strongly reminiscent of that existing between the common school and academy in earlier nineteenth century America.

The pupils in these schools were primarily the children of Irish immigrants who, it is reasonable to suppose, could not keep their children in school as long as other city residents (primarily natives in 1880). Even if enrollment in parochial schools involved a self-selection process, it is surely doubtful that these schools produced as many graduates as the public schools. Moreover, Catholic newspapers began listing grammar school graduates only in the mid-eighties, and then only occasionally. And the registers of the public high school first record entrants from parochial schools only in the nineties. All of these entrants, moreover, are from only one school. All this suggests a small number of Catholic graduates relative to the public schools.

B. The Proportion of Graduates in the Sampled Age Range

Based on the age distribution of public grammar school graduates in selected years. Assumed to be the same in other years and in Catholic schools. From this distribution the proportion who would be the same age as the age cohort sampled in the census was calculated. For example, in the 1880 period the calculation shown below was made. The age distribution of graduates was assumed to be:

age	13	14	15	16	17	18 and older
%	6	20	33	28	10	3

In 1880, then, individuals who were 11-16 constituted 87% of all the graduates (6+20+33+28). In 1885, individuals who had been 11-16 in 1880 were 15-21. They were assumed to constitute only 41% of the graduates of that year (28+10+3).

Information on age distribution came from the following Providence School Committee Reports:

1880 period--1885 (compared to ages of high school entrants in 1885, and in 1879-81. All three very similar.)

1900 period--1895

1915,
1925 periods-from age-grade reports for September (adjusted to the following June) for 1917-19 and 1924-26 respectively.

C. The Sex Ratio of Graduates

--PUBLIC SCHOOLS

1880: from a count of available names
1900: from sex ratio in the class of 1895
1915, 1925: from the ratio among graduates found in the 1915 samples (46% male: identical to ratio in highest grammar class found in 1908 reported in U.S. Immigration Commission Reports (Washington, D.C., U.S. Government Printing Office, 1911) 133-134.

--CATHOLIC SCHOOLS

1880, 1900: from a count of names in the available 1900 lists
1915, 1925: from a count of names in the available 1915 lists

D. The Impact of Population Turnover

Four kinds of migration are relevant. The total number of graduates estimated for a range of years (e.g., for 1897-1904) includes some who left Providence before the Census date (e.g., some who graduated in 1897-1899). The estimate also includes other graduates who entered Providence after the Census date (e.g., some graduates of 1901-1904). A complementary group of sample members must also be considered. Some moved into the city in time to be enumerated in the Census but after having graduated from grammar school in another town. Others left the city after the enumeration and went on to complete grammar school elsewhere. Essentially, the expected number of graduates should be reduced to take account of the last two. But the amounts may be assumed to be roughly equal. Moreover, the impact of migration is not large because as the length of time between the Census and graduation is increased, the probability of someone in the age cohort graduating at that date is reduced. Thus the span of time within which migration must occur in order to matter is severely limited. I estimated the effect very roughly by assuming in-migration and out-migration to be about 5% per year. I then calculated the appropriate proportions of each year's graduates. Some 6 % are involved in the last three periods, and 12 % in the first (which involves a larger range of years in which sample members might graduate, and thus more time for turnover). The proportions which these graduates constitute among all sample members is, of course, far lower.

In short, the expected number of graduates does not have to be increased to take into account sample members who graduated out of town because it has not been reduced to take into account some Providence graduates who could have fallen into the sample. The balance may only be rough but the affected proportions are very low.

It might be noted parenthetically that a school record is in fact available for a sample member who completes grammar school in another city but goes on to high school in Providence. And some record of schooling, albeit incomplete, is available in the last two periods for the individual who entered grammar school in Providence but completed it elsewhere.

Finally, in Table IV-4, a small reduction was made in the expected number of sample members to have attended certain high schools. The expected number, it will be recalled, came from the names in the school records. Some individuals whose names are in the register must have left Providence before the Census year or entered it afterwards. The reduction was based on a conservative estimate of 3% outmigration, 4% inmigration each year, and on an estimate of the time between the census and the average date of high school entry for each list of high school students. The total reduction amounted to 5%-15% in the different schools.

The Expected Number of Entrants Who Fail to Graduate (1915 and 1925 only)

--PUBLIC SCHOOLS

The data collected on sample members include both the name of the last grammar school and the name of the previous school attended by those who reached high school. For this group, the previous school should be the same as the grammar school. The directions to aides were explicit that the latter was only to be copied from high school records and the former was never to be copied from those records.

Thus the number entering high school for whom the previous school is a public grammar school but the last grammar school is not recorded is a measure of the number of missing grammar school records. The proportions missing are : 1915 boys, 11%; 1915 girls, 8.5%; 1925 boys, 5%. The proportions were assumed to hold for all who entered the grammar schools.

--CATHOLIC SCHOOLS

Estimated by two ratios (roughly similar results):

Catholic school non-graduates =

$$1)\quad \left(\text{estimated number of Catholic grammar school graduates}\right) \times \left(\frac{\text{estimated number of public school non-graduates}}{\text{estimated number of public school graduates}}\right)$$

$$2)\quad \left(\text{estimated number of public grammar school graduates}\right) \times \left(\frac{\text{Catholic school enrollment}}{\text{public school enrollment}}\right)$$

The second ratio is derived from the school census figures (Providence School Committee Reports, 1912-1930).

V. OBTAINING INFORMATION ON THE SAMPLE MEMBER'S CAREER

A. TRACING MALE SAMPLE MEMBERS ACROSS TIME TO THE CITY DIRECTORY AND TAX BOOK

The nature of the directories and tax books was discussed above in Section IIA. Tracing sample members across time involves the serious additional problem of correctly matching them to directory listings of a decade or two later. This problem is discussed in detail in a separate paper.

However, two points should be briefly noted here. The first is the possibility of name changes. One frequent kind of change is the Anglicizing of non-English first names (Luigi to Louis) by the individual or by a Census taker or a school teacher. The research assistants who did the tracing were alerted to this problem. Also small changes in spelling - changes which did not substantially affect the sound of a name - Di Meo to De Meo, Brown to Browne - were routinely checked. But other changes usually were not. The dropping of a suffix or prefix from a name and similar changes would have made the pool of similar sounding names too large thus probably increasing rather than decreasing error, and it would have made the rules for tracing too unsystematic and too time-consuming.

The second issue is the effect of having to omit individuals with common names as well as outmigrants from the sample of those who could be successfully traced across time. Omitting those with common names diminishes sample size by about 50% to 33% depending on the period. And sample attrition due to outmigration is also about 45% - 55% in a decade. Thus one can get a feel for the devastating effects of the combination of these factors:

1) A sample of 1000 is

2) Reduced to 500-670 by the common name problem

3) Reduced to 225-370 by outmigration in the first decade

4) Reduced to 100-205 by outmigration in second decade.

Before the reader writes off the results as worthless, he should recall that all historical studies of social structure have had to face these reductions (or the biases involved in limiting the sample to those with common names). The economic attainment data in this study fit into the context of a fair number of roughly comparable studies, but include information on the importance of education for the population that can be studied over time - those who stayed in one city and had relatively uncommon names.

Nevertheless, it will be clear why I have tried to find a way, discussed in a separate paper, of including some information on individuals with common names. It will also be clear why I have traced sample members not found in Providence to the city directories of some surrounding towns. In both cases those included will help raise the sample size, producing more stable results, and the sample will also be more representative.

The towns to which I traced individuals are (a) all the towns contiguous to Providence (Central Falls, Pawtucket, East Providence, Johnston, North Providence and also Cranston from 1910 on) and (b) Warwick in the period in which it was important as a suburb of Providence (from 1920 on). The choice of towns to include was determined in part by practical considerations of available directories and available time and effort for the work. But besides the criterion of contiguity and the obvious need to include Warwick in the later years, the choice was also guided by Sidney Goldstein's study of suburbanization in Rhode Island history. He defines a first ring of suburbs for the period 1900-1930, and a second ring for the period after 1930. Only two towns included in his first ring which I have excluded (Barrington and Smithfield). They are both small.[1]

Finally, in connection with occupational attainment, it should be noted that tracing sample members 10 and 20 years forward in time involves tracing the 1915 and 1925 sample members into the depths of the Depression (in the 20 and 10 year traces respectively). In the latter case, the twenty year trace may serve as something of a corrective. In both cases, but particularly in the former, those unambiguously found in the 1935 trace could be traced to the 1940 directory as well if the direction of the study warrants it and time allows.[2]

B. OBTAINING INFORMATION ON WOMEN SAMPLE MEMBERS FROM THE MARRIAGE RECORDS

For at least a century and a quarter Rhode Island state law has defined the information required on marriage licenses. An Intention to Marry form is first to be filled out by the couple and a copy kept at the city or town hall in which they apply. The person officiating at the marriage is required to sign the form and return it to the town hall within a specified period or face fines. From the returned forms, the city prepared a number of copies of a book listing information on the marriage and on the bride and groom. In using this book, I copied the material in the same form as it appears, with trivial modifications (see the second page of the form for data collection, the first column).[3.]

I have no way to assess whether all marriages were in fact reported in Providence. There is every reason to assume that the overwhelming majority were - but there is no proof. I am, on the other hand, confident that no portion of the book in which marriages were recorded was lost since it is arranged chronologically and has no gaps.

I used marriage license records in two ways. First, the women sample members of 1880 and 1915 were traced to them. And second, a cohort of white-collar brides born in the same years as the 1915 Census sample was selected from these records.

1. TRACING CENSUS AND HIGH SCHOOL SAMPLE MEMBERS TO THE MARRIAGE RECORDS

The key to the use of the marriage records was the convenient indices prepared by the city registrar's office. These indices, which list maiden name of bride, name of groom, date of marriage and page on which recorded, each covered a full decade until 1930, and a quenquenium thereafter. I was thus able to specify an age range in which it was probable that my sample member would marry, and then check all dates involved by looking in a small number of indices.[4] Once again, I have no way to know how well the indices cover the marriages actually recorded in the book. However, the city registrar indicated in the prefaces to the volumes that they were carefully prepared and checked. And it may indicate something of their quality that the references for those included were nearly always accurate, the wrong marriage date or page number appearing perhaps 1% of the time.

The sample member was considered successfully matched if the date of birth was within three years of what it was in the Census (or high school) records, the guardian's name was the same and there was no serious contradiction on other information (race and place of birth). When one such bride was found she was assumed to be the correct match.[5]

One important problem involving the use of the marriage records involved the bride's occupation in the 1880 period. Before 1917, brides' occupations were not indicated on the marriage licenses. "Mobility" reflected in these records can only refer to mobility by marriage. I attempted to trace those brides that were found in the marriage records to the City Directory of the year preceeding their marriages, but very few were found.[6]

A second problem, of course, is that the marriage records provide occupational information only on those who married. On the whole this would be an insignificant bias since the overwhelming majority of women did marry. But those that did not may well have been disproportionately white collar or professional workers. Possible biases will have to be kept in mind.

Finally, the marriage records only include Providence residents. No effort was made to trace women beyond the city limits.

2. COLLECTING A SAMPLE OF BRIDES IN WHITE COLLAR OCCUPATIONS

The marriage license sample was selected by (1) identifying a range of years in which a very high proportion of the 1915 age cohort would have married and (2) sampling all pages in marriage record books for these years (every third page was chosen - i.e. some marriages from every few days during these years were examined); (3) checking for brides with white collar occupations on these pages; (4) checking whether their dates of birth fell within the age range that would have made them 12-15 in March, 1915 when the census was taken; (5) checking whether they were either a) born in Providence or b) living in Providence at the time of the marriage. If all these criteria were met, they were selected.

For the definition of white collar occupations the aides collecting the data were provided with a copy of the Census Bureau's detailed list of occupational categories. Included were those not listed in the three major manual groups (skilled, semiskilled and unskilled) or in the "service worker" groups. Supervisors and proprietors were also included. Aides referred to me all occupations they were not sure were white collar (these were very few) and I of course checked to be sure they were correctly categorizing the ones they were sure about.

VI. THE CLASSIFICATION AND CODING OF OCCUPATIONS

The classification of occupations into major groups which are meant to represent broad, vertically ordered social strata is a problematic endeavor. In the first place it is not clear how well occupations, however cleverly arranged, reflect the most crucial dimensions of social structure. Contemporary studies indicate that they are far from perfect as a proxy for income. And some historical data point in the same direction. For other dimensions of social standing, such as political power, they presumably serve even less well.

But even if one nevertheless chooses to exploit occupation as the best available measure of economic and social position he is confronted with the problem of occupational classification. Assigning occupations to broad groupings is difficult. The group to which a specific occupation should be assigned, the most meaningful broad groupings to use, the social "distance" between groupings, and even their rank order are often unclear.

These problems confront students of contemporary society as well but they are greater in historical research both because they have simply been studied less to date and because less information is available on how well other important characteristics (income, education, prestige, and so forth) correlate with a specific grouping of occupations in various periods. Several articles have argued that the occupational groupings used in studies of contemporary society cannot be mechanically applied to earlier periods, and some attempts to construct alternatives have been made.

The optimal course would be to experiment with a number of classification schemes to see if the choice affected the analysis much. However, to do so presupposes having coded occupations in sufficient detail so that they can be regrouped with as much flexibility as possible. In coding occupations I have tried to retain this flexibility, but at the same time to create a coding system which already presumed a tentative ordering of occupations for preliminary use. That ordering is based on the work of Alba Edwards, and the various modifications of his scheme introduced by the Census Bureau, Stephan Thernstrom, and O.D. Duncan.[1]

It is worth distinguishing three levels of generality in occupational classification: a) the description reported to the census-taker (there are tens of thousands of such descriptions in every federal census); b) a list of several hundred categories of occupations used for detailed tabular presentation (Census Bureau coders convert (a) into (b) by means of the Alphabetical Index of Occupations which has been published in conjunction with each census since 1910); a similar broad categorization is used to classify industries; and c) about a dozen broad groupings to which the detailed categories can be reduced for analysis of major issues.

In the 1920's and 1930's Alba Edwards created a set of about a dozen vertically ordered broad groupings or strata, that would be useful descriptions of socioeconomic position (professionals, clerical workers, semiskilled workers, and so forth). He also proposed a revised detailed list of occupational categories for use by the Census Bureau. His detailed list was more sensitive to vertical ordering of occupations, and fit more neatly into his broad strata. His scheme was based on his knowledge of occupations derived from several decades of work with the Census Bureau. It did not involve the systematic application of criteria such as data on prestige scores, average income, or education. But Edwards makes it clear that he was thinking of each of these, and more particularly of the requisite skills required for the work,as the criteria for vertical classification. The Census Bureau modified his scheme in several ways, notably by actually making it somewhat less clearly vertical - several broad groupings of service workers were added to his list of strata. These service groups are actually descriptions of the job sector and include a wide vertical range. Also, the Bureau was dissatisfied with several categories in Edwards's detailed list. He had included an enormous number of jobs in residual categories such as "operative, not elsewhere classified" and "laborer, not elsewhere classified." The Bureau subdivided these larger categories by industrial sector. Finally it assigned many of the categories on Edwards's detailed list to a different stratum than Edwards had. For example, many categories were shifted from the unskilled to the semi-skilled stratum.

This Census Bureau scheme was used in the 1940 Census and in each census thereafter, with some modifications introduced each time. And for each Census, the Bureau has published an Alphabetical Index to Occupations which lists each functional title in alphabetical order in one list, and each industrial sector description in another, indicating the Census occupational and industrial categories. Thus a specific occupational description is classified into Census Bureau occupational and detailed industrial categories with the help of the index. And then these categories in turn are categorized into broad socio-economic strata.

Thernstrom modified the census scheme first by reconverting the occupations which the Census Bureau had placed in the anomalous service groupings back into the appropriate vertical category, and also by subdividing the white collar workers into high and low. The clerical and sales workers were clearly low, professionals clearly high. Remaining were "managers, officials and proprietors." Those that unambiguously referred to individuals who were relatively wealthy ("banker, manufacturer") were classified as high white collar. The rest were classified low white collar subject to revision later if they possessed a certain amount of assessed property (and if tax data was available).

Duncan's classification scheme, widely used by sociologists, assigns to each of the detailed occupational categories in the 1950 census an exact score. The score is based on the educational attainment and income level typical of individuals in that category. These criteria were used because they had been found to predict well the prestige value of occupations.

The key to the use of either of these classification systems is the use of the census detailed list of occupational categories. I therefore reviewed all the forms of sample members that had been collected in the first year and a half of the project. Each time I encountered a new occupational description I noted it. I then looked up each of these occupational descriptions in the 1950 U.S. Census Alphabetical Index of Occupations.[2] Each was assigned a three-digit code. The first digit represented the broad socio-economic stratum (e.g. unskilled). The combination of first and second represented the particular functional title (e.g. laborer). The three together represented the functional title and industrial sector (e.g. laborer- wire works).[3] Where Thernstrom's list of the occupations in each stratum seemed to contradict the Census Bureau one, or when the meaning of an occupation in Providence at the turn of the century could be expected to differ from the most probable meaning in a national census at mid-century, I made adjustments. But on the whole the assignment of strata based on the Alphabetical Index was a fairly straightforward task.

The occupational description and my code were then typed onto a computer card along with the Census Bureau's codes for the relevant occupational and industrial categories and with the appropriate Duncan score. The resultant set of computer cards was alphabetized and in that form constituted the occupational codebook. The aides who coded occupations with its help were therefore not required to make decisions about occupational classification but rather only to look up an occupation in the code book and indicate on a special form if it did not appear there. This procedure probably kept coding errors at a minimum. The codebook was also useful later as part of a computer program which allowed for easy transition from my three-digit code to the Census and Duncan scores.

Two types of problems arose in using this system. First of all several occupations were difficult to treat in the manner described. Some functional titles clearly have different meanings in different contexts: engineer and fireman are two examples. These were given distinct two-digit codes, one for the unspecified form (a guess had to be made as to which meaning was most likely intended in the majority of unspecified cases), and one for the alternative specified meaning. Thus engineer, unspecified was assumed to be a semiskilled stationary engineer. Engineer, building was assigned the same two-digit code but a different third digit. However, Mechanical Engineer, assumed to be a professional, received a different two-digit code.

Other workers were described only as "works in...," "works at...," "works for...." These also received a separate code for each modifier that followed. A comparison of Census occupations of this type with directory occupations may prove helpful. But tentatively these individuals were classified into the semi-skilled laborer category , unless office work was clearly implied (in which case they were classified as low white collar workers).

Finally still other individuals' "occupations" - particularly in the directory - consisted only of a place or product. For example, "shoe store," or "shoes." I tentatively assumed these individuals owned the place or product in question (unless a large firm was involved, in which case it was easy to determine ownership). However I coded these cases distinctly, so that, for example, it is possible to distinguish "shoe store" from "owner-shoe store" or from "shoemaker - own shop."

Each of these types of occupational descriptions was, however, finally included in the occupational codebook. The second type of problem concerned occupations which had not been included. Some I had accidentally skipped in preparing the codebook, others turned up in later traces or in new samples. When such occupations turned up with any frequency, they were assigned a code and entered into a revised version of the codebook. But it would have been exceedingly tedious to follow this procedure for the many occupational descriptions that occurred only once or twice. Furthermore most such descriptions involved familiar functional titles in a new combination with an industry. Thus electrician had of course been listed in the codebook, but electrician, City Hall had never turned up before. In such cases the unique code I had assigned to that functional title was used and the Census Bureau industry code. In other cases, of course, the functional title itself had not appeared in my codebook. Then the Census Bureau occupation and industry codes were both used. But the appropriate first digit of my code, which assigns the individual to a given socio-economic stratum, was also employed.

Finally, a few occupational descriptions were rare and could not be found in the 1950 Alphabetical Index. I also checked the 1940 index, and the first edition (1937) of the Dictionary of Occupational Titles of the Department of Labor. When a title was listed in one of these two sources it was usually simple to determine what its 1950 census code would have been. But in other cases, I guessed at the appropriate strata and left both the second and third digits blank. Well below 1% of the occupations were classified in this way, and as they are coded in a distinct way they can be omitted or automatically recorded to "unknown" if this seems warranted.

In sum, of about 14,000 occupations coded in the six random samples (four samples of men, two of women), 97% of all functional titles were coded distinctly.

About half of the occupational descriptions included an industrial modification; 80% of these modifications were coded distinctly. Nearly all of the rest of the functional titles and the industrial modifications were coded at the level of detail retained by the Census Bureau. A tiny proportion (.47%) not listed in the Bureau's Index were coded by (assumed) stratum only.

VII. THE ORGANIZATION OF THE DATA COLLECTION AND PROCESSING

The Data Collection. It will be obvious that the data collection and coding described above required a very considerable amount of work. I studied the various sources, considered how best to collect and trace samples, drew up directions for the use of the sources and forms on which to record the information collected, pilot-tested them, and supervised all stages of the work. But the actual selection of 12,000 sample members, and the tracing of each of them to numerous other sources would have been impossible without the help of research assistants. A sample of 12,000 will seem large by contemporary social scientific standards. Recent historical studies have, however, often used samples this large and larger by far. They are easier to collect than most survey data since interviews and mailings are virtually never involved. But in the present case the sample size does not convey the true scope of the effort, since once the samples were collected they were traced by hand to many other sources. Each of those traces, on average, took nearly as long as the original process of selecting the sample member. And, on average, each sample member was traced to 6-8 sources. One may liken the effort, then, to the collection of a census sample of some 80,000-100,000 - except that the "start up" effort of mastering new sources and working with new directions was much greater. The aides who did this work were nearly all undergraduates at Brandeis, Brown, Harvard and M.I.T. (a few were graduate students or recent graduates). In all some 40 individuals have been involved in collecting and processing the material. They worked either full time during the summers or part time during the academic year. A substantial proportion - about half - were partly supported through the Work Study program - a federal program which heavily subsidizes student wages. The others were fully paid by grant funds.

I interviewed and checked references on each applicant, and explained the project in considerable detail to each person hired. Each time an individual began a new operation, involving the mastery of a new set of directions, I worked individually with him for some time. We would work on twenty or thirty cases together and then I would check each of the next fifty or so cases after the aide had completed them. After satisfying myself that the directions would be followed closely, I continued to check a sample of the cases. When I found errors occurring with any frequency, the material was reworked. The sample I chose to check varied in size with the complexity of the task assigned, and with the experience of the aides. I was most compulsive about the forward traces to directories, checking approximately every 10th case throughout. For other operations I generally checked one in fifteen or one in twenty-five. For the simplest operations - such as traces to the tax books or marriage license books or coding non-problem occupations - I checked only about one case in forty.

In all I wrote about 125 single-spaced typewritten pages of directions for the project. But the directions for each particular task were generally 3-6 pages or so, and only once as long as 12 pages. They were not, therefore, impossible for bright, careful undergraduates to master.

Processing the Data. Keypunching and verifying were done in part by agencies specializing in this work and in part (at a fraction of the agency cost) by work study students with prior keypunching experience. The latter appear to have been at least as accurate as the former. This is not surprising since they were told a good deal about the project. They therefore knew what they were keypunching. In weekly half-hour meetings they raised questions with me, and at other times had access to another student aide who knew about all aspects of the work.

Once the data were keypunched and verified a considerable effort was made to carefully check it for errors due to either keypunching or coding. Programs were written to check both for unacceptable values in every variable, and to check for values which were logically inconsistent with values coded in related variables - or were at any rate improbable and worth double checking. A printout of such cases was compared to the forms from which the data had been keypunched. When the form was itself found to contain inaccurate or incomplete data, it was taken again to the primary source.

In addition to checking and correcting the material, the original data set was transformed in a number of other ways and then prepared for processing by the SPSS package of programs which allow for statistical techniques to be used with considerable ease. This preliminary computer work - the work required between the time the data had been keypunched and the time computer runs could be generated for analysis - was great. About 100 computer runs were required - excluding the innumerable ones needed to merely "debug" imperfect programs.[1]

I had no background whatever in computer programming when I began the project, and although I have learned some since I have relied primarily on expert help. In addition to a professional programmer who has worked with the project throughout on a consulting basis, I also involved several undergraduates (majors in Electrical Engineering, and Applied Mathematics and Computer Sciences). With each of these programmers I have worked very closely, discussing programming options, and inspecting test data.

FOOTNOTES

I. Overview: the Design

1. The second cohort was selected from the 1915 rather than the 1925 census since the marital age of the comparable 1925 group falls in the years of the Great Depression whose effects are best avoided if only one later sample is to be compared to the 1880.

III. Selecting the Samples

1. Peter R. Knights, The Plain People of Boston 1830-1860: A Study in City Growth (New York: Oxford University Press, 1971), Appendix C. See also John B. Sharpless and Ray M. Shortridge, "Biased Underenumeration in Census Manuscripts: Methodological Implications," Journal of Urban History, I, #4 (August, 1975), 409-39. The authors suggest that since the proportions of age heaping among native whites, blacks and immigrants were in about the same ratios to each other in 1880-1900 and in 1930-1950, perhaps the ratios of underenumeration are also in the same ratios. If so, there would be roughly a 10% net underenumeration of black males and 20% net underenumeration of foreign born males relative to native white males. An analysis of the trace of high school sample members to the census manuscripts (discussed below) may eventually provide a bit more information on the subject.

2. On the other hand, when the handwriting in the census manuscripts was illegible, the case was skipped, and the first person in the age range listed on the following page(s) was selected instead. Replacements were rarely needed, except in the 1900 Census. The manuscript of that year had been through the smoke and water of a fire and was also difficult to read because a census bureau code had been written over many names.

3. The 1925 Census manuscript listed 35 individuals to a page.

4. The chief drawback of any sample not based on a purely random selection is that it may involve some unsuspected bias. The census pages are arranged by enumeration districts and within district by whatever route the census takers chose to follow. It is difficult to see what socially important variables would be consistently underrepresented by selecting every Nth page of the manuscript.

Nevertheless, if I were to repeat the sampling process I would have selected the pages to sample with the aid of a random number table. It would be a simple matter to determine the number of cases desired, the number of enumeration districts found in the city, and the highest page number in each district. One would then determine the total number of census schedule pages, and, with the help of a random number table, select numbers no greater than this total. As many such numbers would be selected as sample members desired, since the ratio of expected sample numbers to a page is 1:1. If estimates had suggested a ratio of 2:1 one would select half as many pages (more could be selected later, if the estimated ratio proved to be too high; if it proved too low, the last pages selected could be omitted). The system could be adapted to the selection of households by substituting the word "household" for "page" in the preceding discussion. The system could be adapted to the selection of individuals rather than pages by multiplying the total number of pages by the number of lines per page (e.g. 50 in the 1880 Census), and choosing random numbers accordingly. Each resulting random number would be divided by 50 to obtain the page number; the remainder in the division would be the line number. The procedure is quite simple and would add only a few hours to even the most rapid alternative procedures.

I doubt that such a procedure would yield a sample which was any more representative than the one I have in fact selected. But I would have employed it (had I thought of it at the time rather than two years later) and I would recommend it to others. The unsuspected biases that have crept into social science surveys are legendary. R. Christian Johnson suggests a similar procedure. He also argues that a systematic sample excludes neighbors by definition and that since many social phenomena are concentrated geographically, a systematic sample may underrepresent them. He offers no evidence that a 1/50 or 1/100 sampling ratio are small enough so that they would in fact lead to such biases, only that they might be. In any case, the sampling ratios for pages in the present study were much lower (men: 1/4, 1/6, 1/8 and 1/8 in successive periods;women: 1/6 and 1/8). These biases, if they in fact exist at all, should be very low. One could also argue that their effect is probably counteracted by sampling whole pages. R. Christian Johnson, "A Procedure for Sampling the Manuscript Census Schedules," Journal of Interdisciplinary History, VIII (winter, 1978), 515-530.

5. It might also be noted that by including individuals in a given age range on sampled pages, I had to include either more or fewer pairs of siblings than a pure random sample would have included. That is, the two choices are either to include all siblings in the age range or only one individual per family. The same problem would have arisen with a pure systematic, or a truly random sample of households. The sample may be thought of as comparable to individuals in a given age range from randomly selected families.

Finally, within institutions for children every Nth inmate (every 4th in 1880, every 6th in 1900 and so forth) was selected. The institutionalized children are therefore represented in correct proportion.

6. The oversamples include all Jewish boys and 3/4 of the Jewish girls; half the Italian and Irish boys and 3/8 of the girls. In each case the proportions were obtained by sampling certain pages, in the manner described above.

7. Rhode Island Commissioner of Labor, "The 1905 Rhode Island Census" in Twentieth Report of the Commissioner of Labor (Providence: State Printing Office, 1907),
On using "Russian" as a surrogate for East European Jew see, for example, Marshall Sklare, America's Jews, (New York: Random House, 1971), 72.

8. The Census takers were instructed to distinguish Russia as a place of birth from Poland (then part of Russia, Austria and Germany). And apparently they did so consistently, since a great many individuals are in fact listed as having been born in Poland. Also the 1905 state Census takers clearly had already followed the same practice, because had they not done so the proportion of Jews among Russians would not have been 95%. Polish Catholics listed as Russians would have driven the proportion down.

9. This approach is meant to demonstrate that sampling variability has not produced an unusually unrepresentative sample. Suppose one obtained a sample in which the mean of a given variable was more unlike the population mean than the means of 95% of the mathematically possible samples. The chances of then drawing a second such sample would be 25/10,000 ($^5/_{100}$ X $^5/_{100}$).

On the other hand if a bias were inherent in the sampling procedure, this test would not uncover it since both samples would share the bias. As I indicated above, I see no reason to suppose such a bias.

10. Some data was not available in all four periods. The most glaring omission is occupational data in the Census of 1925. That Census alone also did not provide data on unemployment nor on the parental grandfather's place of birth. School attendance, literacy and the marital status of parents in single parent households was unavailable in both the 1915 and 1925 state censuses. The employment status (employer or employee) of the head was available only in 1915, and the information in Part Ic of the form, as indicated there, was available only in 1900. Ward number is not provided in the 1880 census, but it can be derived from enumeration district numbers with the help of maps and descriptions provided by the Cartographic Division, National Archives.

11. One dimension it excludes, however, is whether siblings are working or in school. Since all siblings 13-16 (or 12-15) are included in the sample, the question really refers only to those older than 16 (or 15) who are still living at home. The code indicating the presence of such siblings in the household might serve at least to isolate those families in which this factor might arise.

12. The number of girls entering some high schools was much larger than the number of boys. Consequently only entrants in the Census year, or a systematic sample of them, were selected. Naturally, the numbers will be weighted in these cases.

13. In selecting the entrants, I could have made the high school samples more comparable to the Census sample by selecting an age cohort (rather than entire entering classes). That is, I could have included all those entrants 13-16 in 1880, 12-15 in 1879, and so forth. Such a cohort would differ little, I suspect, from the one I did select, which included all entrants in a given year or range of years regardless of age. It seemed worth suffering a small loss in comparability of samples to get a picture of an entire class.

III. Obtaining Additional Information on Family Background

1. Guardians in the 1925 sample who were not found in the city or in the house directories of that year were also traced to the city directory of the preceding and succeeding year, because the 1925 Census did not include a question on occupation. Only from these traces, then, would such information be available. The same procedure was used for the 1880 period to compensate in part for the absence of the house directory. But relatively few cases were found in this way. Comparisons in the paper on the directories are almost exclusively within a period, so this difference in technique by period - even if it produced greater differences in results than I estimate - is largely irrelevant to those results.

2. Coarser residential information from the Census is machine readable, namely the enumeration district numbers.

3. The directory occasionally lists, instead of an occupation, a firm name in italics and bracketed by parentheses after the individual's name. In all such cases the information was copied as it appeared, but the firm name was also searched in the directory. In virtually every case it was found, followed by one or more names in parentheses, including the name of the individual in whom I was interested. This form of listing appears to indicate that the individual is an owner or part-owner of the firm. In some cases the names in parentheses are followed by titles: president, treasurer, etc. The individual is thus clearly a company officer. All this was copied and coded with the occupation.

4. The last two digits of the value were truncated because they were usually zeros. When an individual shared property with others (except his wife), he was assumed to hold an equal share, and the amount was divided by the number of owners listed. When an individual had been found in the directory to own a firm, the firm's name was also looked up, and its value, if indicated, was added to the amounts listed for the individual. If he was only part owner of the firm, only part of its value was assigned to him (once again, the total value was divided by the number of owners). Corporation officers were generally not assigned the value of their corporation. The exception was businesses that were clearly family firms. Individuals were not assigned the value listed by their names if they were labelled as trustees or executors. Finally it should be noted that when two owners are listed for one property they are listed only once. Since the tax book is organized in alphabetical order, it will generally be possible to catch the listing only if the individual being traced is the first owner listed. Presumably the omissions are random (among partners).

5. Information on family structure also could have been collected, but it would have greatly increased the time needed for the work. This information could be retrieved at a later date if necessary because the location of the sample members in the Census manuscript was recorded.

IV. Obtaining Information on School Attainment

1. See the discussions in Historical Statistics of the United States: Colonial Times to 1970 (Washington, D.C.: U.S. Government Printing Office, 1975), Pt. 7, 360-367; John K. Folger and Charles B. Nam, Education of the American Population (Washington, D. C.: U.S. Government Printing Office, 1967), Chapter 1 and appendix. On state censuses see Henry J. Dubester, State Censuses: an Annotated Bibliography of Census of Population Taken after 1790 by States and Territories of the United States (Washington, D.C.: U.S. Government Printing Office, 1948; prepared for the U.S. Census Bureau).

2. On Worcester: The Bulletin of the Worcester Historical Society, October, 1977. On Beverley and Somerville: Michael B. Katz, The Irony of Early School Reform: Educational Innovation in Mid-Nineteenth Century Massachusetts (Cambridge, Mass.: Harvard University Press, 1968), On Iowa, Richard J. Jensen and Mark Friedberger, "Education and Social Structure: An Historical Study of Iowa 1870-1930" (Final Report: National

Institute of Education, 1976). Information on the whereabouts of New York City school records is available from the New York City Board of Education (unfortunately they are not kept centrally). Information on Salem, the Massachusetts reform schools, Pittsburgh, Quincy and Boston wasprovided by Carl Kaestle (Dept. of History, University of Wisconsin, Madison), Barbara Brenzel, (Graduate School of Education, Harvard), Carolynne Schumacher (Dept. of History,University of Pittsburgh) · Warren Leon (Dept. of History, Harvard University) Carol Srole (Dept. of History, U.C.L.A.) respectively.

3. Only one high school's cards made some attempt to preserve the notion of academic class as opposed to courses completed each calendar year. Possibly the other high schools that used cards only used the credit system.

4. The records of the public high schools were kept in the following form:

	1880	1900	1915	1925
Providence High School	register	X	X	X
Classical High School	X	register	register	register
English High School	X	register	register to 1915; cards thereafter	record cards
Manual Training (Technical) High School	X	record cards	record cards	record cards
Hope High School	X	record cards	record cards	record cards

Unfortunately the three sets of record cards are not interfiled. The English High School records are in one file, the Hope records of 1900 and 1915 in another. All other high school records including those of the evening high school are interfiled with the grammar school record cards. Hence tracing sample members to these records (described below) meant three separate operations.

5. One kind of document I chose not to use were statements of tuition bills paid by pupils. The two Catholic high schools have preserved such lists but they have numerous drawbacks. Both schools for example operated elementary as well as secondary schools, and tuition payments might well have been for the former (amounts seemed to prove an inconclusive guide).

6. However, one grammar school's records have been preserved since the 1890's. They appear on large sheets of paper which have conveniently been folded to fit in the same large file as the other grammar school records.

7. Sometimes it was also possible to include background information on the graduates that would be useful for record linkage (such as address, and guardian's name taken from the lists of graduates prepared by the schools or from teacher registers).

8. The evolution of the laws pertaining to the working papers and Census are described in Charles Carrol, Public Education in Rhode Island (Providence: E.L. Freeman, 1918), 200-202, passim.

9. Some attempt to gain a better picture of the grammar school graduates will also be made by tracing the graduates of 1881 to the high school records to determine who continued and who did not and then to trace the names to the 1880 Census to obtain information on their socioeconomic origins. Since there are only 240 of these students the task should not prove too time consuming.

10. In the 1900 period the first two groups will have to be lumped together. There is no point in selecting a sample of the especially poorly schooled both because even fewer than in 1880 received so little education and because compulsory education laws enacted in the eighties were being enforced by this time. Hence families who did give their children such a short education had an incentive to lie to the census taker about it. In the two later periods information on those not graduating from grammar school is available in the public school records and in any case far more did graduate from grammar school by then, making distinctions in the group that did not less important.

11. Finally, sample members have already been traced to the catalogues of two institutions which cannot be neatly categorized as "secondary" or "post-secondary". These are the Rhode Island School of Design and the Rhode Island Normal School (later the Rhode Island College of Education). In each case, annual catalogues were available which included lists of students. Many colleges also listed students names in their annual catalogues or in an alumni directory. Since the number attending college is small, and since many of these attended a few local colleges (notably Brown which did publish such a directory) it may also be possible to obtain information on college graduation with relative ease.

12. On the Commercial Schools see Edmund J. James, "Commercial Education" in Nicholas M. Butler, Education in the United States: A Series of Monographs (Albany, N.Y.: J.B. Lyons, 1900), 653-704.

13. My procedure was similar to the procedure discussed in an altogether different context by Michael B. Katz and John Tiller, "Record Linkage for Everyman: A Semi-Automated Process," Historical Methods Newsletter, 5(September, 1972), 144-150. The Philadelphia Social History Project's linkage program is described in Theodore Hershberg, Alan Burstein and Robert Dickhorn, "Record Linkage," Historical Methods Newsletter, 9 (March/June, 1976), 137-163.

14. Had the computer's search turned up many names that were really the same but might not appear so in a hand trace (because they were spelled differently), retaining the use of the soundexing program would have been important. But in fact there were virtually no such names. On the contrary, the machine was producing many pairs of cases in which the soundex code was the same but the names were obviously different. using the linkage program rather than only the soundex program would presumably have resulted in the rejection of these pairs. But since I could not exploit that program, I had to reject a large number of such pairs by hand.

15. If a sample member was linked to a graduate who had the same name but who was apparently living in another part of the city, the city directory was consulted to check whether his family had perhaps moved between the census and graduation dates. I did not use the same technique with Catholic or private schools, since geographic boundaries were usually not applicable, or at any rate less rigid.

16. In the case of the linkages to private school records I also used the information in the 1880 and 1900 census on whether the sample member was in school, working, or neither to determine whether the two names referred to one person).

17. Several points are not obvious from the data collection form:

a. From the date left last high school or grammar ashool, and age, it is possible to calculate school leaving age.

b. Curriculum information is very often missing for high school students. This is partly because within certain high schools (e.g. Classical in 1900, or English and Science Department, 1880) further subdivision by curriculum was unimportant. And partly it is because "curriculum" was apparently often defined in terms of course credits at graduation. Thus in some schools curriculum data is available for graduates, but not for non-graduates.

c. The highest semester graded was distinguished from time in school. The former reflects academic progress.

d. No single course was chosen to reflect academic achievement since required courses varied widely over time and between institutions. This is even true of the significance of a course such as "English." A range of grades (every other one) was therefore selected. The grades were selected from the end of the first year of attendance at high school - before too many students had dropped out but after an initial period of adjustment. For those who had dropped out already, semester or quarter grades were taken. (If a student repeated a grade, the first set of marks was used.) A similar procedure was used in recording grammar school grades: those from the beginning of the sixth grade were taken. Finally, for those remaining in high school at least 5 semesters, the next to last semester's marks were taken for comparison with the first year marks.

e. I also recorded the source from which the informantion was copied on the last grammar, secondary and post-secondary enrollment. When the information came from a mere list of names (implying a less certain linkage) the strength of the match (see linkage section above) was also coded.

f. Very few records were available for students who quit while still in primary school, but primary schooling was often indicated on records of those continuing onto grammar school. Probably the primary school records were transferred to grammar schools with the students. And when the records were centralized it appears that only those in grammar schools and high schools were the focus of the effort.

g. It seemed too cumbersome to record each shift in pre-secondary enrollment, particularly given the fact that such information often was not available because sources other than school records had been used. But it is worth noting in passing that were such information complete, it might well have been interesting to examine the number of schools different children attended. Given the geographic mobility of the population, there must have been a lot of movement from school to school. And given the possibility of greater geographic mobility among the lower social classes it is possible that one reason for differences in educational attainment by social groups could be related to the effects of this instability.

h. The sources from the two institutions that could not be classified neatly as secondary or post-secondary - the Rhode Island School of Design and the Rhode Island Normal School - were tentatively treated in the following manner. If other school records indicate that the sample member completed high school, this schooling was treated as post-secondary. Otherwise it was recorded as secondary.

18. Business schools and all postsecondary institutions are also relevant in this context. But I have discussed these omissions above. My concern here is with the regular elementary and secondary grades.

19. Providence School Committee, Reports for 1875-1930 (Providence: City Printers); the latter has the title Directory of the Public Schools in some years; Sadlier's Catholic Directory, Almanac and Ordo (New York, 1881); published annually: title and publisher vary; Americao Lapati "A History of Catholic Education in Rhode Island" (unpublished Ph.D. dissertation, Department of Education, Boston College, 1958); Mary C. Morgan, Little Sketch of the Work of the Sisters of Mercy in Providence, R.I. from 1851 to 1893 (Providence: J.A. Reid and R. A. Reid, 1893), Charles Carroll, Public Education in Rhode Island (Providence: E.L. Freeman, 1918), and Rhode Island: Three Centuries of Democracy,(New York: Lewis Historical Publishing Co., 1932); Thomas B. Stockwell, A History of Public Education in Rhode Island from 1636 to 1876 (Providence: City Printer, 1876); Rhode Island Commissioner of Education, Annual Report for 1892-1921. Inquiries concerning unpublished tabulations at the Providence School Committee, the state office of education and the Catholic school office (as well as at city, state and diocesan record centers) turned up almost no additional information.

20 Statistical tests comparing expected and actual frequencies, such as those used in assessing the adequacy of the Census sample in Section II above, have not been employed here. The distortions in the school data could of course occur from sampling bias. But it is more likely that they occur from records that were missed due to problems in the sources and problems in tracing. Hence in almost every case the expected frequency is larger than the actual frequency, rather than sometimes larger, sometimes smaller as sampling variability would probably create (in an unbiased sample). Also, the expected values themselves are based on rough estimates and much of the deviation of actual from expected values could in fact be caused by errors in these estimates. The statistical tests were nevertheless made. They provide a sense of the magnitude of the deviation of actual from expected values. These deviations are thus compared to those that would be considered acceptable in sampling. The deviations in Table IV-2 are indeed statistically significant in most cases, although not in all.

It is also possible that those not found in school records were missed because their names were hopelessly misspelled by the census taker, or miscopied by the aides collecting the data. If so, however, their guardians also would probably not have been found when traced to the city directory of the census year. For it is the last name, not the first name, which is surely most likely to be misspelled. Yet the great majority of these guardians (63% to

84% in different samples) were in fact found in the directory of the census year. Another indication that name problems are not the cause of misclassification may be seen by comparing the Italians of 1915 and 1925. In 1915 they are three times as numerous among those for whom no record was found as they are among the others. This result might be caused by their attrition or by misspelled names. But if it is caused by the latter, the results should be the same for the 1925 sample. Yet in fact in that sample the proportion of Italians in the group not found in the school records is almost identical to their proportion among the sample members that were found there.

21. The entering class in the selected public high schools includes parochial and private grammar school graduates as well as public school graduates. The group is not, therefore, comparable to the public grammar school graduates in Table IV-2. However, it seems true (as such a comparison would suggest) that very high proportions of the grammar school graduates continued on to high school. The Providence School Committee Annual Report for 1889 includes (pp. 16-17) a comparison of pupils in the highest grammar school grade and entrants to high school who came from these schools in the succeeding year. The comparison covers the classes entering high school in 1885-88. I assumed, unrealistically, that fully 95% of those entering the highest grammar grade completed it. The proportion of grammar school graduates continuing on to high school would then be 62%. And in the age grade reports of the period 1917-27, the ratio of public high school freshmen to those in the highest grammar school grade averaged 87%. It is inflated, of course, by parochial and private grammar school graduates entering the public high school, and perhaps by more repeaters in the higher grade. But even allowing for these factors the enrollment ratios suggest that those graduating from public grammar schools--(probably four-fifths of all graduates) must still have been going on to high school in large numbers.

22. On the whole, the individuals for whom school information is lacking are more likely to have fathers in low manual occupations, and fathers who own less property. And they are more likely to be of "new immigrant" stock. These relationships are all predictable. But if anything, they seem rather weaker than what might have been expected.

23. Estimating continuation rates by comparing the numbers in successive grades (e.g.: the number in grade 2 divided by the number in grade 1) leads to enormous errors, at least for the lower grades. For one thing, large numbers of pupils were regularly held back in grade 1, thus making comparison of the enrollment in grades 1 and 2 meaningless. A similar pattern seems to have been typical in the first grammar school grade (grade 5). Consequently, enrollment in that grade is usually more than 100 percent of the enrollment in grade 4. The continuation rates between grade 2 and grade 3 and between grade 3 and grade 4 are consistent in age grade reports available for 1917-1927. In each case, they are about .9. On the basis of these continuation rates one would have to conclude that no less than 30% to 40% of the students left school before reaching the grammar school grades (10 % dropping out in grade 2, another 10% in grade 3 and at least 5% both in grade 1 and in grade 4). The discrepancy between this result and those obtained by focusing on the grade level of 13 year olds (discussed in the text) is large. Also in the sample data, records are actually available (no estimates involved) to show that 70% to 82% of the

individuals in different samples from 1915 and 1925 did reach grammar school. Any missing records would mean the true proportion is actually even higher. Consequently, no more than 18% to 21% could conceivably have received less than a fifth grade education.

It is not entirely clear why the calculations of continuation ratios from enrollment by grade are misleading. Different promotion rates for different grades may have affected enrollment. More important appears to be the fact that some children were channeled into ungraded classrooms of various kinds (e.g.: for disciplinary cases). Students were taken out of the regular grade sequence and enrolled in these ungraded rooms, mostly between the ages of 9 and 12. Thus in 1925 the number in grades 2, 3 and 4 was 11,625, of which 5,425 were 9-12 year olds. But another 1,493 9-12 year olds were in the ungraded classes. Virtually all had been put there after they had turned 8, most when they were 9 or older. The timing of their removal from the regular grade sequence apparently accounts for much of the seemingly high drop-out rate in grades 1-4.

24. The age grade reports include public school students only, the sample data includes all students. But the public schools constituted 80% - 85% of total enrollment in any given year. This proportion was remarkably stable for each age (figures by age from the school census are reported in the Providence School Committee, Reports each year). It is extremely unlikely that the enrollment pattern in the Catholic schools (where virtually all of the nonpublic enrollment was to be found) would be so different that it could affect the totals for the entire age cohort by an appreciable amount.

25. I suspect that the estimate for 1915 girls not reaching grammar school is a bit low. It results partly from what seems to me an unlikely circumstance, which is indicated in Table IV-2 and also affects Tables IV-5 and IV-6. The estimates there imply that the records of many more public grammar school graduates were missed in the sample of girls than in the sample of boys (i.e.: the difference between expected and actual is larger in the sample of girls). Since the two samples involve identical records used in identical ways, such a result seems unlikely.

V. Obtaining Information on the Sample Members' Careers

1. Sidney Goldstein and Kurt B. Mayer, Metropolitanization and Population Change in Rhode Island (Providence, R.I.: Rhode Island Development Council, 1962), 27-30, passim.

It should be noted that these towns are not all "suburbs." In some cases, notably much of Pawtucket and Central Falls, they are simply extensions of the same urban industrial "inner city" area as Providence.

Tracing those not found in the Providence directory to the directories of surrounding towns involved a considerable effort. Generally, each town had its own directories. Tracing an individual to 5 towns meant 5 traces. A further complication was that the larger the area within which the sample member is sought, the greater the chance that two listings will be found with the sample member's name, or that even if only one is found he will not in fact be the sample member. To reduce these dangers somewhat all the towns outside Providence were treated as one unit. Even if a sample member was found in one of these towns, he was traced to all the others. If his name turned up in a second town the result of the trace was coded as ambiguous.

In tracing to the tax books of surrounding towns and cities I had to rely on an incomplete set of books. When none was available, I labelled the case with a special code.

2. A somewhat similar problem arose in connection with the 11 and 12-year-olds collected in the 1880 samples (regular, Black and unschooled). To trace them to the 1890 directory, like the 13-16-year-olds, seemed unwise, since they would be rather younger than the others. At the outset of a career each year might be important in assessing occupational attainment. Also, they might even have been young enough to have been omitted by the directory altogether. Therefore they were traced to the directory of 1892, rather than to that of 1890. For comparability, a full decade was left before the second decennial trace: they were traced to the 1902 directory rather than the 1900 one.

3. Historical Records Survey (Division of Women's and Professional Projects, Works Progress Administration), "Summary of Legislation Concerning Vital Statistics in Rhode Island" (mimeograph: 1937). Available in the Brown University Library.

4. I chose 1881-1910 for the 1880 sample members (they would have been 41-46 in 1910) and 1916-1945 for the 1915 sample members (they would have been 42-45 in 1945). Some women may have first married after age 42-46, the last age represented in the indices I used. This number is likely to be very small indeed. But certain groups of women may have been more likely to be among the older brides. Furthermore, I divided the dates into a more probable and a less probable category of years (the last 10 years in each range being the less probable ones). Names in the more probable range of ages were checked in the records first. Finally, the very few names which appeared a large number of times in the indices were excluded altogether as too difficult to trace. They are, however, identifiable by a distinct code.

5. When none of the brides met these criteria for identification other rules had to be employed (some 10% of the cases which were coded distinctly). Most of these cases involved high school sample members for whom only one guardian's name was known or orphans in institutions for whom no guardian's name was known. The children were identified if the name was an uncommon one (found only once in the marriage license indices) and the date of birth was within one year of the Census (or high school) record one. Another type of ambiguous case occurred rarely but often enough to be noticeable: 1%-3% of the brides clearly lied about their age. When the names of both guardians were identical, and all other available information matched, and the birthdate was within 8 years of the Census one, the match was considered correct. In still other cases, the match was considered too ambiguous to resolve (about 1% of the 1915 and 6% of the 1880 sample).

6. It may prove fairly simple to trace these brides to the 1885 Rhode Island State Census which is alphabetically arranged by ward, if the direction of the analysis warrants it.

VI. The Classification and Coding of Occupations

1. This discussion is based on Alba Edwards, "A Social Economic Grouping of the Gainful Workers of the United States," Journal of the American Statistical Association 27 (1933), 377-387, and Comparative Occupational Statistics for the United States, 1870-1940 (Washington, D.C., U.S. Government Printing Office, 1943); James Scoville, "The Development and Relevance of U.S. Occupational Data", Industrial and Labor Relations Review, 19 (1965) 71-79; Gladys Palmer, "The Convertibility List of Occupations and the Problems of Developing it", Journal of the American Statistical Association 34 (1939) 693-708; Otis Dudley Duncan, "A Socioeconomic Index for All Occupations" in Albert J. Reiss, Occupations and Social Status (New York: Free Press, 1961); Stephan Thernstrom, The Other Bostonians: Poverty and Progress in the American Metropolis, 1880-1970 (Cambridge, Mass: Harvard University Press, 1973), Appendix B; Theodore Hershberg, Michael Katz, Stuart Blummen, Laurence Glasco, Clyde Guffen, "Occupation and Ethnicity in Five Nineteenth Century Cities: A Collaborative Inquiry," Historical Methods Newsletter, 7 (June, 1974), 174-217.

Christopher Jencks, et al, Inequality: A Reassessment of the Effects of Family and Schooling in America (New York: Harper Colophon Books, 1973) 176-208; Michael B. Katz, The People of Hamilton, Ontario West 48-93; Donald J. Treiman, "A Standard Occupational Prestige Scale for Use with Historical Data" The Journal of Interdisciplinary History VII, 2 (Autumn, 1976) 283-305; Robert W. Hodge, Paul M. Siegel and Peter H. Rossi, "Occupational Prestige in the United States 1925-1963," American Journal of Sociology, 70 (November, 1964) 286-302.

2. Since both Duncan and Thernstrom relied on the 1950 Alphabetical Index that seemed the logical one for me to use also. The only earlier one incorporating Edwards' work is from 1940 and is very similar to the 1950 one. Also, in 1940 a category semi-professional workers was used, and in 1950 all these workers were "promoted" to professionals. I retained the semi-professional classification, and followed Thernstrom in classifying these workers as low white collar.

3. The simplest solution would have been to create separate codebooks for functional titles and for industries. But to do so would have meant two separate coding operations, and in any case I had allowed only three computer card columns for coding occupations (too few for two detailed classification systems). So I chose what was a more difficult way to construct a codebook but an easier one for the aides to use: each combination of functional title and industrial sector was listed. Thus laborer was listed about a hundred times, (laborer, ice factory; laborer, wire works; laborer, not specified, etc.). The aides, then, simply looked up the occupational description in the occupational codebook as it had been recorded on the form and copied the appropriate three digit code. Since a large number of codes were needed, an alphanumeric scheme was used.

If more than the available number of third digit characters is needed, a range of second digit characters is used for one occupation and the third digits can then be repeated. For example Z1A=laborer, ice factory; Z3A=laborer, wireworks.

VII: The Organization of the Data Collection and Processing

1. Some examples of data transformations, other than checking and correcting the material follow. 1) Each occupation was assigned the appropriate U.S. Census occupational and industrial codes as well as the appropriate Duncan score. 2) When two or more sample members had been selected from one household family background data had been recorded only on the eldest. It was now copied onto the records of other siblings. 3) The ethnic oversamples were completed by duplicating certain cases from the random sample (see page11). 4) the automated linkage to school records used for some samples was undertaken (see page 43). 5) The creation of analytically meaningful variables for use in the SPSS system were created from the categories of material collected from the sources. For example, "Ethnicity" was defined as race for blacks, place of birth of the guardians of whites. Similarly "Educational attainment," "Family Class" and others were created.

ID number:	1-7	
Surname:	8-21	
Forename:	22-37	

I. CENSUS DATA

Part A. Sample Member

Enumeration dist. or Book	38-40	
Page	41-43	
Ward	44-45	
Address ________________		
Sex	46	
Race	47	
Age	48-49	
Place of birth ________	50-51	
Occupation ________________	52-54	
Unemployed?	55	
At school?	56	
Literacy	57	

Part B. Family Background

Child/Guardian Relationship	58-59	
Single parent household?	60	
FATHER		
Name ________________		
Age	61-62	
Place of birth ________	63-64	
His father's pl.b ________	65-66	
Occupation ________________	67-69	
Unemployed?	70	
Literacy	71	
BLANK	72-79	
Card number	80	1

1925: ADDITIONAL INFORMATION

IQ ⟶

App./Court./Init./Relia./Coop.→

Add. Ed. ________________

Occ. Pref. ________________

Placement ________________
(firm, salary, position)

____

ID number:	1-7	
MOTHER		
Name ________________		
Age	8-9	
Place of birth ________	10-11	
Her father's pl.b ________	12-13	
Occupation ________________	14-16	
Unemployed?	17	
Literacy	18	
Head of family: first initial	19	
citizen?	20	
(1915 only) employER/EE?	21	
SIBLINGS		
Older males	22	
Older females	23	
Younger males	24	
Younger females	25	
If sample member foreign born, oldest U.S. born sibling	26-27	
If sample member non-R.I. born, oldest R.I. born sibling	28-29	
KIN OTHER THAN PARENTS OR SIBLINGS		
Parents (in-law) of head of household	30	
Other, 1 gen. older than hd.	31	
Same generation as head	32	
1 gener. younger than head	33	
2 gener. younger than head	34	
2 gener. older than head	3	
LODGERS	36	
SERVANTS	37	

Part C. 1900 Census only

Time in U.S.: Child	38-39	
Father	40-41	
Mother	42-43	
Mother: Total # living Chdn.	44-45	
House owned?	46	

II. RELATED INFORMATION

Father (hd. of F.) in direc.?	47	
Occupation ________________	48-50	
Religion (source: ________)	51	

III. PROPERTY TAX

Father (hd. of F.) in tax book?	52	
Real	53-55	
Personal	56-58	
BLANK	59-62	0
	63-67	
	68-79	
Card number	80	2

Note: Columns labelled 'blank' were later used for supplemental data

IV. SCHOOL DATA

ID number 1-7

Part A. General

Date of birth 8-11
Place of birth__________ 12-14
Address____________________
Parent's name (ini.)________ 15
Occupation__________________
______________________ 16-18
Father's nativity 19-20
Mother's nativity 21-22
Cause left last school 23

Part B. High School

Date entered FIRST H.S. 24-27
Date left LAST H.S. 28-31
Drop/reenter?
1) Semester left 32
For how many months 33
2) Semester left 34
For how many months? 35
First high school (name) 36-38
Age at entry 39-42
Previous school 43-45
Curriculum 46-47
High school diploma? 48
Highest semester graded 49
For transfers only:
Last high school(name) 50-52
Curriculum 53-54
(diploma? enter above)
Highest semester graded 55-56
Other high school? 57-59
Next to last semester's grades
(for those graded 5 semesters or more)
1) Subject Category 60
Grade 61
2) Subject Category 62
Grade 63
3) Subject Category 64
Grade 65
4) Subject Category 66
Grade 67
5) Subject Category 68
Grade 69
6) Subject Category 70
Grade 71
Deportment Grade (1880) 72
Whole days absent 73-74
BLANK 75-79
Card number 80 3

ID number 1-7
First year's grades:
1) Subject Category 8
Grade 9
2) Subject Category 10
Grade 11
3) Subject Category 12
Grade 13
4) Subject Category 14
Grade 15
5) Subject Category 16
Grade 17
6) Subject Category 18
Grade 19
Deportment Grade (1880) 20
Whole days absent 21-22

Part C. Grammar School

Name of last 23-25
Source 26
Highest semester graded 27-28
Graduate? 29
Date left 30-33
Grammar school grades:
1) Subject Category 34
Grade 35
2) Subject Category 36
Grade 37
3) Subject Category 38
Grade 39
4) Subject Category 40
Grade 41
5) Subject Category 42
Grade 43
6) Subject Category 44
Grade 45
Effort Grade 46
Deportment Grade 47
Whole days absent 48-49

V. TEN YEAR TRACE

Source 50

Part A. Occupation

____
______________________ 51-53
Address____________________

Part B. Property Tax

Found in tax book? 54
Real 55-57
Personal 58-60

VI. TWENTY YEAR TRACE

Source 61

Part A. Occupation

____
______________________ 62-64
Address____________________

Part B. Property Tax

Found in tax book? 65
Real 66-68
Personal 69-71
BLANK 72-79
Card number 80 4

ID Number: 1-7 ☐ ☐ ☐☐☐☐ ☐

Initials 8-9 ☐☐

Surname____________________________

Forename___________________________

VII. TEN YEAR OUT-OF-TOWN TRACE

Source 10-11 ☐ ☐

City or Town 12 ☐

Occupation_________________________

____

_____________________________13-15 ☐☐☐

Address___________________________

Found in tax book? 16 ☐

Real 17-19 ☐☐☐

Personal 20-22 ☐☐☐

VIII.TWENTY YEAR OUT-OF-TOWN TRACE

Source 23 ☐

City or Town 24 ☐

Occupation_________________________

____

_____________________________25-27 ☐☐☐

Address___________________________

Found in tax book? 28 ☐

Real 29-31 ☐☐☐

Personal:1(Tang. or unspec.) 32-34 ☐☐☐

Personal:2(intang.) 35-37 ☐☐☐

IX. THIRD FORWARD TRACE

NOT USED

Source 38 ☐

Occupation_________________________

____

_____________________________39-41 ☐☐☐

Address___________________________

Found in tax book? 42 ☐

Real 43-45 ☐☐☐

Personal:1(Tang. or unspec.) 46-48 ☐☐☐

Personal:2(intang.) 49-51 ☐☐☐

X. THIRD OUT-OF-TOWN TRACE

Source 52 ☐

City or Town 53 ☐

Occupation_________________________

____

_____________________________54-56 ☐☐☐

Address___________________________

Found in tax book? 57 ☐

Real 58-60 ☐☐☐

Personal:1(Tang. or unspec.) 61-63 ☐☐☐

Personal:2(intang.) 64-66 ☐☐☐

Supplemental occup. codes 67-70 ☐☐☐☐

BLANK 71-79 ☐☐☐☐☐☐☐☐☐

Card Number 80 5

ID # 1-7 ☐ ☐ ☐☐☐☐ - ☐

Initials 8-9 ☐☐

XI. MARRIAGE INFORMATION

Traces to ML Index: first 10 ☐

supplementary 11 ☐

Source 12-13 ☐ ☐

Surname 14-22 (for ML sample) ☐☐☐☐☐☐☐☐☐

Forename 23-31 ☐☐☐☐☐☐☐☐☐

Date of marriage 32-35 ☐☐☐☐

Pl. of Mar. if not Prov.______

SAMPLE MEMBER

Date of birth____/______

Residence______ 36 ☐

Age 37-38 ☐☐

Color 39 ☐

Occupation______

______

______ 40-42 ☐☐☐

Birthplace______ 43-45 ☐☐☐

Birthplace of father______ 46-47 ☐☐

Birthplace of mother______ 48-49 ☐☐

Father's name______

Mother's name______

No. of marriage 50 ☐

Occupation of father______

______

______ 51-53 ☐☐☐

SPOUSE

Full name______

Date of birth____/______

Residence______ 54 ☐

Age 55-56 ☐☐

Occupation______

______

______ 57-59 ☐☐☐

Birthplace______ 60-62 ☐☐☐

Birthplace of father______ 63-64 ☐☐

Birthplace of mother______ 65-66 ☐☐

Father's name______

Mother's name______

No. of marriage 67 ☐

Occupation of father______

______

______ 68-70 ☐☐☐

Did EITHER sign with mark? 71 ☐

Denomination of marriage 72 ☐

______

Supplemental occup. Codes 73-76 ☐☐☐☐

BLANK 77-79 ☐☐☐

Card number 80 6

Surname______

Forename______

XII. TRACES TO M.L. INDEX

First traces:

Date	Volume	Page	Contr. Initial
___/___/___			
___/___/___			
___/___/___			

Supplemental traces

___/___/___			
___/___/___			
___/___/___			

ID No. 1-7 ☐ ☐ ☐☐☐☐ ☐

Initials 8-9 ☐☐

XIII. TRACES BASED ON MARRIAGE LICENSES

NOT USED

Father of sample member

Found in directory? 10-11 ☐☐

Found in tax book? 12 ☐

Total property value 13-15 ☐☐☐

Father of spouse

Found in directory? 16-17 ☐☐

Found in tax book? 18 ☐

Total property value 19-21 ☐☐☐

Sample member found in directory 22-23 ☐☐

Found in tax book? 24 ☐

Total property value 25-27 ☐☐☐

Spouse found in directory 28-29 ☐☐

Found in tax book? 30 ☐

Total property value 31-33 ☐☐☐

XIV. 1935 CENSUS TRACE

Found in Census? 34-35 ☐☐

Census tract number 36-38 ☐☐☐

St. address______

If hd. of family: # in fam 39-40 ☐☐

Marital status 41 ☐

Read or write English 42 ☐

Usually work? 43 ☐

If no, seeking work? 44 ☐

Usual Occupation______

______

______ 45-47 ☐☐☐

Class of worker 48 ☐

Present occupation______

______

______ 49-51 ☐☐☐

Class of worker 52 ☐

If not at work, why not? 53 ☐

If not at work, # of mos. idle 54-55 ☐☐

months employed 1935 56-57 ☐☐

Supplemental occup'tn codes 58-61 ☐☐☐☐

BLANK 62-70 ☐☐☐☐☐☐☐☐☐

71-79 ☐☐☐☐☐☐☐☐☐

Card number 80 7